DOCUMENT RESUME

ED 030 156 24 CG 004 123

By-Treffinger, Donald J., Ripple, Richard E.

The Effects of Programed Instruction in Productive Thinking on Verbal Creativity and Problem Solving Among Elementary School Pupils. Final Report.

Cornell Univ., Ithaca, N.Y.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No-BR-8-B-002
Pub Date Dec 68
Grant-OEG-0-8-080002-0220-010
Note-238p.
EDRS Price MF-$1.00 HC-$12.00
Descriptors-*Creativity, Elementary Schools, Elementary School Students, Instructional Innovation, *Problem Solving, *Productive Thinking, *Programed Instruction

The effectiveness of the Productive Thinking Program in developing creative thinking and problem solving abilities among 370 pupils in grades four through seven was examined. An exploratory investigation of the program's instructional content was also undertaken. At all four grade levels studied, instructed pupils' scores on Covington's Childhood Attitude Inventory for Problem Solving were significantly greater than those of controls'. There were no significant indications of transfer from the instructional program to scores on verbal tests of creative thinking, or problem solving batteries, at any of the grade levels tested. Pupils written responses were evaluated in five general categories: Memory-Organization, Production, Reorganization, Judgment-Evaluation, and Attitude. Those high and low on each of these factors were compared with a sub-sample of control pupils on each of three creative problem solving criteria, with no significant differences among these groups for any response factor across all grade levels. Results were interpreted in terms of three factors: conditions under which the program was administered, criterion difficulty, and complexity of the abilities involved. (Author/CJ)

ED030156

BR-8-B-002
PA-24
OE-BR

BR 8B002
PA 24

FINAL REPORT
Project No. 8-B-002
Grant No. OEG-0-8-08002-0220-010

THE EFFECTS OF PROGRAMED INSTRUCTION IN PRODUCTIVE THINKING ON VERBAL CREATIVITY AND PROBLEM SOLVING AMONG ELEMENTARY SCHOOL PUPILS

December 1968

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FINAL REPORT
Project Number 8-B-002
Grant Number OEG-0-8-080002-0220-010

THE EFFECTS OF PROGRAMED INSTRUCTION IN PRODUCTIVE THINKING ON VERBAL CREATIVITY AND PROBLEM SOLVING AMONG ELEMENTARY SCHOOL PUPILS

Donald J. Treffinger
Purdue University
W. Lafayette, Indiana

and

Richard E. Ripple
Cornell University
Ithaca, New York

December 1968

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
Office of Education
Bureau of Research

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

TABLE OF CONTENTS

Chapter Page

Table of Contents . . . ii
List of Tables . . . iv

I. INTRODUCTION . . . 1
  Summary . . . 1
  Introduction . . . 4
  General Statement of the Problem . . . 8
  Specific Statement of the Problem . . . 10
  Previous Research with the Productive Thinking Program . . . 11

II. METHODS AND PROCEDURES . . . 18
  Selection and Assignment of Pupils . . . 18
  The Experimental Instructional Materials . . . 26
  Procedures . . . 30
  Measuring Instruments . . . 31
    Tests of Verbal Creativity . . . 31
    Lorge Thorndike Intelligence Test . . . 44
    Arithmetic Pre-Test . . . 45
    General Problem Solving Test . . . 47
    Arithmetic Puzzles Test . . . 56
    Arithmetic Problems - Text Form . . . 58
    Make Up Problems Test . . . 60
    Childhood Attitude Inventory for Problem Solving . . . 62
  Treatment of the Data . . . 64
    Principal Instructional-Control Comparisons 64
    Internal Analyses . . . 70
  Relation to the Specific Statement of the Problem . . . 76
  Chapter Summary . . . 79

III. RESULTS OF THE STUDY . . . 82
  Pre-Test Comparisons . . . 82
  Instructional-Control Comparisons . . . 87
    Verbal Creativity . . . 88
    General Problem Solving . . . 98
    Arithmetic Problem Solving: Text Form . . . 105
    Arithmetic Problem Solving: Puzzle Form . . . 105
    Make Up Problems . . . 110
    Attitude Inventory . . . 111
  Internal Analyses . . . 113
    Correlational Data . . . 114
    Analyses of Covariance . . . 117
  Chapter Summary . . . 126

Table of Contents (Continued)

Chapter Page

IV. DISCUSSION AND CONCLUSIONS 128
  Effects of Instruction: Instructional-Control Comparisons 128
    Verbal Creativity 128
    General Problem Solving 129
    Arithmetic Problem Solving: Text Form 130
    Arithmetic Problem Solving: Puzzle Form . 131
    Make Up Problems 131
    Attitude Inventory 132
    Summary and Interpretation 132
  Comparisons By Pupil Program Response Factors 144
  Conclusions 147
  Implications and Suggestions for Research . 151

REFERENCES CITED 157

APPENDICES

A. Arithmetic Pre-test (With Levels of Item Difficulty and Discrimination) 164
B. General Problem Solving Test 171
C. Arithmetic Problem Solving Test - Text Form (With Levels of Item Discrimination and Difficulty) 181
D. Arithmetic Problem Solving Test - Puzzle Form (With Levels of Item Discrimination and Difficulty) 190
E. Make Up Problems Test 197
F. Childhood Attitude Inventory for Problem Solving 204
G. Intercorrelations of Problem Solving Tests, IQ, Arithmetic Pre-Test, and Creativity Total Score 211
H. Summary of Categories Assigned to Pupils' Program Responses 213
I. Teacher Rating Procedure and Creativity Definition 219
J. Sample Pages from the Productive Thinking Program 225

LIST OF TABLES

Table Page

1 Participating Schools 19

2 Lorge Thorndike Intelligence Test Scores (By Classes) 21

3 Distribution of Parents' Occupational Levels by Grade and Class 22

4 Summary of Pupils Reading One Year or More Below Grade Level (By Grade and Class) 23

5 Summary of Sex Distributions by Grade and Class 24

6 Distribution of Pupils By Sex and Treatment Groups 25

7 Lorge Thorndike Verbal Intelligence Scores (By Grade Level and Condition) 25

8 Test-Retest Correlations: Torrance Tests of Creative Thinking (Control Pupils Only) 35

9 Inter Scorer Reliability: Torrance Tests of Verbal Creativity 36

10 Coefficients of Stability of Scoring for Originality Sub-Tests 37

11 Correlations Between Teachers' Rankings of Pupil Creativity and Pupils' Standardized Total Creativity Pre-Test Scores 40

12 Arithmetic Problem Solving Test (Puzzle Form) Placement of Problems By Grade Level 57

13 Summary of Reliability Coefficients for Instruments Used in This Study 65

14 Summary of Statistical Analyses Conducted at Each Grade Level 80

15 Analysis of Variance: Pre-Test Variables (Grade Four) 83

16 Analysis of Variance: Pre-Test Variables (Grade Five) 84

LIST OF TABLES (Continued)

Table | Page

17 Analysis of Variance: Pre-Test Variables (Grade Six) 85

18 Analysis of Variance: Pre-Test Variables (Grade Seven) 86

19a Analysis of Covariance Comparison of Instructional and Control Creativity Post-Tests, Creativity Pre-Test Covaried (Grade Four) 89

19b Analysis of Covariance Comparison of Instructional and Control Creativity Post-Tests, Creativity Pre-Test Covaried (Grade Five) 90

19c Analysis of Covariance Comparison of Instructional and Control Creativity Post-Tests, Creativity Pre-Test Covaried (Grade Six) 91

19d Analysis of Covariance Comparison of Instructional and Control Creativity Post-Tests, Creativity Pre-Test Covaried (Grade Seven) 92

20a Analysis of Covariance Comparison of Instructional and Control Creativity Post-Tests, Creativity Pre-Test and IQ Covaried (Grade Four) 93

20b Analysis of Covariance Comparison of Instructional and Control Creativity Post-Tests, Creativity Pre-Test and IQ Covaried (Grade Five) 94

20c Analysis of Covariance Comparison of Instructional and Control Creativity Post-Tests, Creativity Pre-Test and IQ Covaried (Grade Six) 95

20d Analysis of Covariance Comparison of Instructional and Control Creativity Post-Tests, Creativity Pre-Test and IQ Covaried (Grade Seven) 96

LIST OF TABLES (Continued)

Table		Page
21	General Problem Solving, Type C Problems; Proportion of Solutions by Grade and Condition	101
22	General Problem Solving, Type C Problems; Analysis of Covariance: Total Number of Solutions, IQ covaried, by Grade and Condition	102
23	General Problem Solving, Insolvable Problem; Proportion of Acceptable Responses, by Grade and Condition	103
24	General Problem Solving, Type O; Summary of Analyses of Covariance By Problems, Controlling for the Effects of IQ, By Grade and Condition	104
25	General Problem Solving, Type O; Analysis of Covariance, Standardized Total Score, IQ Covaried, By Grade and Condition	106
26	Arithmetic Problem Solving: Text Form Analysis of Covariance, Controlling for Pre-Test and IQ, By Grade and Condition	107
27	Arithmetic Problem Solving: Puzzle Form Proportion of Pupils Solving Each Problem By Grade and Condition	108
28	Arithmetic Problem Solving: Puzzle Form Analysis of Covariance on Total Score (Controlling for IQ and Pre-Test) By Grade and Condition	110
29	Make Up Problems Test Comparisons By Grade and Condition	111
30	Pupil Attitude Inventory Comparisons By Grade, Part of Inventory and Condition	112
31	Pearson Product Moment Correlations and Fisher Z-Coefficients Between Program Response Factors, and Creativity and Problem Solving Criteria, By Grade Level	116

LIST OF TABLES (Continued)

Table		Page
32	Analyses of Covariance: Type O Problem Solving Memory-Organization Factor (IQ Covaried)	117
33	Analyses of Covariance: Type C Problem Solving Memory-Organization Factor (IQ Covaried)	118
34	Analyses of Covariance: Post-Creativity Total Score, Memory-Organization Factor (IQ and Pre-Test Covaried)	118
35	Analyses of Covariance: Type O Problem Solving Production Factor (IQ Covaried)	120
36	Analyses of Covariance: Type C Problem Solving Production Factor (IQ Covaried)	120
37	Analyses of Covariance: Post-Creativity Total Score Production Factor (Pre-Test and IQ Covaried)	121
38	Analyses of Covariance: Type O Problem Solving Reorganization Factor (IQ Covaried)	121
39	Analyses of Covariance, Type C Problem Solving, Reorganization Factor (IQ Covaried)	122
40	Analyses of Covariance, Post-Creativity Total Score, Reorganization Factor (IQ and Pre-Tes Covaried)	122
41	Analyses of Covariance, Type O Problem Solving Judgment Factor (IQ Covaried)	123
42	Analyses of Covariance, Type C Problem Solving Judgment Factor (IQ Covaried)	123
43	Analyses of Covariance, Post-Creativity Total Score Judgment Factor (IQ and Pre-Test Covaried)	124
44	Analyses of Covariance, Type O Problem Solving Attitude Factor (IQ Covaried)	124

LIST OF TABLES (Continued)

Table		Page
45	Analyses of Covariance, Type C Problem Solving Attitude Factor (IQ Covaried)	125
46	Analyses of Covariance, Post-Creativity Total Score Attitude Factor (IQ and Pre-Test Covaried)	125

CHAPTER I

INTRODUCTION

Summary

This research investigated the effectiveness of the Productive Thinking Program (Covington, Crutchfield, and Davies, 1966) in developing verbal creativity and problem solving abilities among pupils in grades four through seven. The program, which the writers considered an innovative attempt to provide direct instruction in creative problem solving, was studied in several ways. First, its effects on pupils' verbal creativity scores, when intelligence and pre-test scores were statistically controlled, were examined. Next, the question of transfer from the programed instruction to several problem solving criteria was investigated. These criteria consisted of a General Problem Solving test, including both Type O and Type C problems (following Davis, 1966); two forms of an Arithmetic Problem Solving test, to investigate non-specific transfer of learning to subject-matter problem solving; the Make Up Problems test (Getzels and Jackson, 1962); and, an early form of the Childhood Attitude Inventory for Problem Solving (Covington, 1967). The final specific purpose of the study was to undertake an exploratory investigation of the internal structure or instructional content of the programed materials.

Three hundred seventy pupils enrolled in 16 classes from six public school systems in central New York state participated in the study. There were four classes at each of the four grade levels studied. Two classes at each grade level were randomly assigned to the instructional condition; the remaining classes served as controls. Classes were generally similar with respect to socio-economic status, sex and intelligence distributions, and proportion of pupils with reading deficiencies.

At each of the four grade levels there were significant differences favoring instructed pupils on the Childhood Attitude Inventory for Problem Solving (Part I, general attitudes, and Total Score). For verbal creativity, general problem solving, arithmetic problem solving (both forms), and Make Up Problems test scores, there were no indications of a significant effect of the instructional treatment at any of the four grade levels studied.

In order to investigate the internal structure of the instructional program, pupils' written responses were analyzed. A randomly drawn sub-sample of twenty pupils from the instructional group at each grade level was used for these analyses. Five general factors were identified as components of the 175 responses required of pupils. These were: Memory-Organization; Production; Reorganization; Judgment-Evaluation; and Attitudes. Correlations between pupils' scores on these factors and each of three creative problem solving criteria

(Verbal Creativity Total Score; Total Type O Problem Score; Total Type C Problem Score) were computed. At each grade level, and for each creative problem solving criterion, differences between correlations with the five factors were tested for significance. There were no significant differences for any of the three creative problem solving criteria at any of the four grade levels.

Then, for each of the five factors, and separately at each grade level, two groups were identified: pupils high on the factor, and pupils low on the factor (approximately upper and lower thirds of the sub-sample). For each factor, at each grade level, scores of these pupils and a randomly drawn sub-sample of control pupils were compared on each of the three creative problem solving criteria, controlling for IQ differences among groups. Of sixty analyses, there were two significant F-ratios; these were interpreted as artifacts of the number of analyses conducted.

The results of the study were interpreted in terms of three factors: the conditions under which the instructional materials were used (without teacher participation, one lesson per day on sixteen consecutive days); the difficulty of the criterion measures; and, the likelihood that more extensive training and practice must be offered to develop such complex cognitive abilities.

Introduction

Interest in creativity among psychological and educational researchers in America is quite recent. Before the mid-twentieth Century, very little research in psychology or education was undertaken in the area of creative thinking. Early mental tests seldom included measures of creative thinking abilities, and even among those scholars who were interested in understanding what comprised the nature of intelligence, creativity received little attention (41).

In 1950, J. P. Guilford, then President of the American Psychological Association, delivered an address in which he commented on the lack of research in this area, stated some propositions about possible relationships between creative thinking and human intellectual abilities in general, and described briefly a program of research which he and his colleagues were undertaking to investigate the nature of human intelligence as a multi-dimensional construct (36). Since then, research addressed to creative thinking has increased greatly in volume. Razik (55), for example, has prepared a bibliography of studies of creativity and problem solving which contains more than four thousand entries. That publication has already been supplemented by an additional review covering the period from the publication of Razik's bibliography through 1966 (9).

Guilford himself has contributed, with his associates, to the study of creativity, both in terms of research addressed to creative thinking and problem solving abilities and through the formulation of the Structure of Intellect Model which has shown considerable value as a heuristic device for researchers interested in human intelligence (37, 41).

The increase in volume of research in the area of creative thinking has also been the subject of a recent paper by Guilford (40). He describes the acceleration of interest that has occurred since 1950 and proposes that it will probably continue. He proposes the need for research in two directions: first, toward a greater understanding of the process of creative thinking; and, toward a more complete understanding of the conditions that influence creative thinking.

The great increase in research on creativity has also created serious problems for the scholar. Controversies have raged, and many are yet quite unsettled, concerning the definition of creativity, problems of measurement and assessment, and the feasibility of attempts to foster creativity in educational settings. So much has the term "creativity" been used in the psychological and educational literature, and with such variation in meaning and usage, that many scholars now shun it as "unscientific." Some scholars have been openly antagonistic to attempts at developing creative thinking abilities (2).

As there has been considerable controversy and confusion among scholars concerned with creativity, Davis makes a similar observation about problem-solving research:

"Research in human problem solving has a well earned reputation for being the most chaotic of all identifiable categories of human learning (25, p. 36)." This "chaos" has at least two dimensions: first, the lack of consistent general definition of problem solving among authors; second, the use of a large number of differing tasks as criteria for problem solving performance. These problems, in Davis' view, lead to difficulty in comparing experimental results, and to an inability to make generally valid statements about problem solving.

Many scholars, however, have continued to be fascinated with the investigation of creativity and problem solving. In addition to research directed towards understanding the nature of these abilities, there has been an increasing amount of research directed towards attempting to facilitate them through instruction. Some have attempted to "train" originality through verbal learning experimental procedures (46). Others have developed programs based on the "brainstorming" approach (50, 51). Torrance and his associates have developed workbook materials, audio tapes, and records intended to help children develop creative thinking abilities (23, 47, 70). Covington, Crutchfield, and Davies have developed a series of programed instructional materials which are designed to develop creative problem solving abilities and positive attitudes about creative thinking among elementary school pupils (17).

The importance of continuing research on creativity and problem solving has been stressed in several sources. The present writers, following Guilford (36, 40), Torrance (64, 65, 67, 68), Parnes (52), Covington (13, 15), and others, contend that creativity and problem solving represents an area of critical importance in the psychological study of cognitive processes, as well as an area of research which has important social consequences. Guilford has, in the opinion of these writers, stated the case for the importance of research on creative thinking in a succinct way:

> It is apparent that the solutions to numerous human problems are dependent upon education of the world's population....An informed people...is a creative, problem-solving people. In a real sense, mankind is involved in a race between expanding education on the one hand, and threatened disaster, perhaps oblivion, on the other. (40, p. 12.)

From the point of view of the educator, the problem of facilitating the development of creative thinking and problem solving abilities through instruction is one of the most important in this area of research. Of course, the educator may well also be interested in better understanding the nature of these abilities, how they may be more effectively assessed among pupils, and the relationships among personality characteristics and these abilities. But the question, "What can we actually do to help pupils develop creative thinking and problem solving abilities?" remains critical.

The present study is concerned with the development of creative thinking and problem solving abilities among pupils in grades four through seven, through the use of a set of programed instructional materials, called the Productive Thinking Program (17).

There are several important questions concerning the effectiveness of these materials which will be examined in this study. In addition to questions of the effectiveness of the instructional materials, other concerns are also addressed in this study. The materials are viewed as one approach, from among a number of possibilities, used in order to examine more carefully the general problem of facilitating creative thinking and problem solving abilities. It is necessary, then, to present in detail the problem to which this study addresses itself. The statement of the problem will be presented in two parts: first, an overview, or general statement of the problem; second, a detailed examination of the specific components of the problem.

General Statement of the Problem

This study addresses itself, at the broadest level of description, to the question, "Can creative thinking and problem solving abilities be taught to pupils in grades four through seven by direct instruction?"

The first general aspect of the problem is suggested by the words "direct instruction." By direct instruction is meant an attempt to develop creative thinking and problem solving abilities in which the specific instructional content is creativity and problem solving as such. The Productive

Thinking Program (17) constitutes one attempt at such direct instruction, and its effectiveness will be systematically investigated at four grade levels.

The second general aspect of the problem is comprised of questions dealing with the psychological study of learning and transfer with respect to the abilities and attitudes which the instructional program proposes to develop. The authors of the instructional materials propose that such abilities and attitudes should be generalizable to a wide range of problem solving situations (22). Although the question of transfer has been raised before (10, 56), there is little research evidence to support or deny conclusively Crutchfield and Covington's position that positive transfer should occur. Another psychological issue which is addressed by this study has to do with the question, "What is learned?" Although the programed instructional sequence has been utilized in several previous research projects, this rather important question has not yet been raised. In all of the studies which have been reported previously, it has either been assumed that the pupils learn from the program what its authors propose it teaches, or such considerations have not been made at all. Knowledge of what is learned by pupils who study the Productive Thinking Program, through classification and analysis of pupils' responses, is thus a major aspect of the problem to which this study is addressed.

These general dimensions of the problem can be better understood through a detailed examination of more specific questions. Several such questions have been generated from the general statement of the problem; these will be presented as the specific statement of the problem.

Specific Statement of the Problem

The specific questions to which this study is addressed, which have been developed from the general problem stated above, are classified according to:

A. Those emerging from the first part of the general statement of the problem, dealing specifically with the instructional materials:

1. What abilities are developed in the program?

2. If some abilities are more effectively developed than others by the instructional materials, to what extent do the abilities taught bear differentially on pupils' problem solving performance?

3. At what grade level(s), if at all, may the instructional materials be recommended for classroom use?

4. To what extent, if at all, must the instructional materials be revised, modified, or supplemented by other activities for optimal effectiveness in the classroom?

B. Those emerging from the second part of the general statement of the problem, dealing with the psychological questions of learning and transfer:

5. To what extent, if at all, does the instructional treatment (i.e., studying the Productive Thinking Program) facilitate performance on tests of verbal creativity, when pre-test scores and IQ are statistically controlled?

6. To what extent, if at all, does what is learned or developed in the instructional treatment lead to positive transfer to general problem solving situations?

7. To what extent, if at all, does the learning from the instructional treatment lead to positive transfer to arithmetic problem solving? (It should be noted that arithmetic problem solving has been selected as a form of problem solving representative of problems which pupils must typically confront in the classroom, rather than for any intrinsic relationship it might be thought to have with human problem solving or the specific content of the programed instructional materials.)

Previous Research with the Productive Thinking Program

Covington and Crutchfield reported two studies with fifth-and sixth-grade pupils, in which a preliminary version of the instructional program (comprised of 13 rather than 16 lessons) was used (16).

In the first study, 195 pupils from Berkeley and vicinity participated. These pupils were from four fifth-grade and two sixth-grade classes. Two of the fifth-grade classes and one of the sixth-grade classes were designated as instructional groups; pupils in these classes studied the 13 lesson programed sequence. The instructional materials were used for one hour per day for a three week period. Each child worked individually on each lesson. Control classes received a shorter set of booklets which did not provide instruction in creative problem solving. Following the training period, an eight hour post-test battery was administered to all pupils. A one-hour follow up test battery was also given five months later.

Pupils in both conditions were compared on several measures, including: number of problem-clarifying questions asked, number and rated quality of ideas generated, and number of solutions achieved. For each of these measures, the 98 pupils in the instructional group markedly outperformed uninstructed pupils. Differences between proportions of pupils solving the problems in instructional and in control classes far exceeded statistical significance. On some problems, the instructed pupils outperformed controls by as much as a three to one ratio. Significant differences favoring pupils in the instructed classes were also observed on various tests of divergent thinking abilities.

Follow up testing after five months showed continuing superiority for the instructed children over the control children.

The second study reported by Covington and Crutchfield utilized all 16 lessons which presently comprise the instructional sequence. A total sample of 286 pupils was studied. Results again indicated marked superiority for instructed pupils over controls on several criterion measures. The facilitative effects of the program were stronger at the fifth-grade level than at the sixth-grade level, however, when data were analyzed separately by grade level.

Ripple and Dacey have reported research in which a ten lesson experimental version of the program was used with eighth-grade pupils (24, 56). The total sample in this research consisted of 136 pupils from ten classes. Five classes were randomly assigned to the instructional condition; pupils in these classes studied the instructional program. Pupils in control classes were given only their normal classroom instruction. All pupils in both instructional and control groups were given a post-test battery of divergent thinking measures, which were scored for imagination, fluency, flexibility, and originality. There were no significant differences between instructional and control pupils on any of these measures. Twenty-five pupils in each condition were asked to attempt to solve the Maier two string problem. Although slightly more pupils in the instructed group successfully solved the problem than in the control group, the difference did not reach statistical significance. However, instructed pupils did solve the problem significantly _faster_ than control pupils. Although Ripple and Dacey concluded that their results supported non-specific transfer effects from the training materials to an actual problem solving criterion, their results were, on the whole, considerably less emphatic than those reported in the early studies by Covington and Crutchfield. Ripple and Dacey noted that the Covington and Crutchfield results were much less emphatic at the sixth-grade level than at the fifth, and suggested that it may be the case that more challenging or differently oriented programs are required as grade level increases.

Recent reports from the Berkeley Creativity Project have been of research with the programed instructional materials in several samples, all at the fifth-grade level.

Covington reported data from fifth-grade classes in the San Francisco Bay area (15). The instructed pupils in this study were administered the instructional program over a one month period. The instructional group (N=54) and the control group (N=54) were matched with respect to initial problem solving ability, IQ, and school achievement. Sex and racial distributions were comparable. Several criteria indicated that the post-test performance of the instructed pupils was generally superior to that of the control pupils. Measures included problem booklets with several multiple choice items presented at various stages in the problem sequence, designed to assess pupils' choices of ideas and procedures, given the factual constraints of the problem. Instructed pupils showed consistent superiority over controls on such items. Eleven instructed pupils, compared with only three control pupils, solved the problem during the early stages of presentation. Following the administration of several multiple choice questions, pupils were again directed to propose solutions. At this time, 15 additional pupils in the instructional group solved the problem, compared with five control pupils. Covington observed that, although there was a decided training effect, it was also true that a large number of instructed pupils did not show performance which differed from that of control pupils. He suggested that these results may point to the need for additional supplementary practice, revision of parts of the program, or both.

Olton et al., have reported another study in which the programed materials were used with fifth-grade pupils (49). In this study, 704 pupils from 44 fifth-grade classes in the Racine, Wisconsin public schools were involved. The instructional materials were used in 22 classes, one lesson per day, for four days of each of four school weeks. Students worked individually, at their own pace; the teacher's role was held to a minimum, emphasizing a test of the effectiveness of the materials themselves. Pre-and post-test batteries, including extensive problems thought to emphasize convergent thinking and divergent thinking, brief problems emphasizing convergent and divergent thinking, and a general verbal test intended to assess the proposed "master thinking skill" were administered. Mean performance of the instructed pupils surpassed that of control pupils on 30 of the 40 internal and post-test measures. Eleven of those differences reached statistical significance.

The authors observed that, although significant differences favoring instructed pupils were found, such differences were of considerably smaller absolute magnitude than those indicated in previous studies (16). Failure to find larger differences was interpreted as reflecting differences between the sutdies with respect to teacher involvement, and the pace at which the programed materials were administered to the pupils. In previous studies, classroom teachers were deliberately instructed to supplement the lessons (although that fact was not explicitly stated in the 1965 paper). In addition, the materials had not been used as rapidly (almost one lesson each day) in previous studies. Olton et al., concluded that large differences would not be likely to result from such a severe test of the materials.

In more recent work several modifications in the procedure for administering the programed materials have been made explicit (21). These include:

(1) "spacing" of the administration of the lessons, so that at least one full day intervenes between the presentation of each lesson, and often slowing the pace to as few as two lessons per week over an eight week period;

(2) increasing teacher involvement, utilizing a newly prepared Teacher's Guide (18), including warm up and review discussions and "highlighting" key points of the lessons;

(3) providing pupils with supplementary worksheets, which emphasize the "key point" of each lesson and provide the pupil with additional practice.

Recent research has utilized the programed instructional materials with these modifications. Results suggest that the instructed pupils are substantially superior to control pupils on a number of measures. A recently completed study involved fifth-grade pupils in the San Francisco area. The programed lessons were administered at the rate of two per week, over an eight week period. The program was supplemented by teacher discussions, ten to fifteen minutes in length, in conjunction with each lesson, and by using sixteen supplementary exercises. Instructed pupils were considerably superior to control pupils on most criterion measures. One criterion measure, intended to probe transfer effects to a typical classroom problem, involved writing an essay on "Poverty In Plenty." While instructed pupils' descriptions of poverty did not differ from control pupils' descriptions, instructed pupils made significantly more statements of possible causes of poverty. In addition, instructed pupils gave far greater (3:1 ratio) suggestions for "solutions" to poverty (21).[1]

---

[1] The writers with to thank Dr. Robert Olton, of the Berkeley Creativity Project staff, for providing preliminary information about the results of this research, prior to the publication of the paper cited.

CHAPTER II

METHODS AND PROCEDURES

This Chapter will review the selection and assignment of pupils, the experimental materials, the procedures, and the measuring instruments used in this study. The treatment of the data will be discussed in detail and the reliability and validity of all measuring instruments will also be assessed.

Selection and Assignment of Pupils

This study involved pupils from 16 classes in grades four, five, six, and seven, from six public school systems in central New York State. There were four classes at each grade level. The school systems which participated were selected from approximately 30 public and parochial school systems in the area served by the Finger Lakes Region Supplementary Educational Center, a regional center under Title III of the Elementary and Secondary Education Act of 1965. A summary of the participating classes and school systems, by grade level, is given in Table 1.

Table 1

Participating Schools

System	Location	Grades	No. of Classes
Union Springs	Union Springs, N.Y.	4, 5	4
McGraw Central	McGraw, N. Y.	4, 5, 6	6
Moravia Elemen.	Moravia, N. Y.	6	1
Cincinnatus	Cincinnatus, N. Y.	6	1
Interlaken	Interlaken, N. Y.	7	2
Ovid	Ovid, N. Y.	7	2

Classes at each grade level were selected according to correspondence on the following criteria:

(1) Classes should be generally similar with respect to distribution of verbal intelligence scores, and should not represent "homogeneous" groups;

(2) Classes should reflect generally similar distribution of socio-economic class levels, based on parental occupation data;

(3) Classes should be similar with respect to total numbers of pupils reading one full year or more below grade level;

(4) Classes should be similar with respect to distribution of male and female pupils (that is, similar proportions of boys and girls among classes at any grade level.)

On the basis of these criteria, four classes at each grade level were identified to participate in the study. No classes which satisfactorily met these criteria and were invited to participate in the study declined to be involved. Tables 2 through 7 summarize the comparisons of participating classes with respect to the selection criteria. Table 2 summarizes the mean, standard deviation, and range of each class on the Lorge Thorndike Verbal Intelligence Test (44) and Table 7 summarizes those data by grade level and treatment groups. Table 3 summarizes the distributions of parental occupation levels among all classes, based on the classification developed by Warner and his associates (77). Table 4 summarizes the number of pupils in each class reading one year or more below grade level, on the basis of standardized test data provided by school administrators.[2] Table 5 summarizes the number of pupils in each participating class, by sexes.

Since, except at the sixth-grade level, each school system was represented by two classes, assignment to experimental and control conditions was made randomly within school systems. That is, at each grade level, one class from each of two school systems was randomly selected for the experimental

---

[1] Pupils' scores in grades four, five, and six were based on data from the Iowa Basic Skills tests, given either in May 1967 or September 1967. Seventh-grade classes received the California Reading Test in September 1967.

Table 2

Lorge Thorndike Verbal Intelligence Test Scores[1]

(By Classes)

Class-Grade[2]	N[3]	Mean	S. D.	Range
411	28	109.43	15.19	76-137
421	19	107.16	13.37	84-141
432	28	108.54	17.66	72-147
442	24	103.04	15.68	76-134
511	21	105.71	14.07	81-130
521	23	105.96	13.49	80-138
532	23	109.91	16.49	80-143
542	22	105.00	14.61	74-131
611	26	106.31	15.48	73-136
621	24	108.96	11.38	73-128
632	27	106.11	14.08	87-137
642	20	103.25	15.29	73-132
711	21	108.14	16.08	84-135
721	22	116.00	8.63	101-131
732	18	106.38	16.10	69-129
742	29	114.59	11.25	94-138

---

[1] In grades four through six, Level Three, Form A, Verbal; in grade seven, Level Four, Form A, Verbal.

[2] Identification code numbers for classes: Grade (4-7), class number (1-4); condition (1=instructional, 2=control).

[3] Number of pupils present for testing (prior to pupil elimination based on incomplete pre-test data.)

Table 3

Distribution of Parents' Occupational Levels

By Grade and Class[1]

Grade-Class	No. of Parents in each of Warner's Class Strata N	1	2	3	4	5	6	7	Unclassified[2]
411	31	1	4	3	5	14	3	1	0
421	27	2	4	2	7	6	0	0	6
432	29	1	2	4	4	12	1	0	5
442	27	1	2	4	7	8	2	3	0
511	24	1	4	3	4	12	0	0	0
521	27	0	2	3	1	13	2	1	5
532	24	1	3	0	4	9	5	0	2
542	25	2	2	4	3	12	0	0	2
611	25	0	0	5	5	12	1	0	2
621	28	0	6	1	5	7	1	0	8
632	27	3	1	2	2	16	2	1	0
642	29	3	1	5	4	10	4	1	1
711	27	1	3	3	3	6	2	1	8
721	24	1	6	3	2	4	2	0	6
732	29	3	6	4	0	4	4	2	6
742	24	1	4	5	3	10	1	0	0

---

[1] Based on the classification scheme used by Warner et al., (1960); see p. 139-140.

[2] Unclassified includes pupils for whom data was not available, pupils under institutional or non-parental care, and occupational definitions which were unable to be classified in the strata identified by Warner et al., from the lists provided by the schools

Table 4

Summary of Numbers of Pupils
Reading One Year or More
Below Grade Level (By Grades)[1]

Grade, Class	N	No. (Pct.) 1 year or more Below Grade Level	
411	31	3	(10.0)
421	27	5	(18.5)
432	29	3	(10.3)
442	27	7	(25.9)
511	24	8	(33.3)
521	27	4	(14.8)
532	24	7	(29.2)
542	25	5	(20.0)
611	25	8	(32.0)
621	28	4	(14.3)
632	27	5	(18.5)
642	29	6	(20.7)
711	27	0	(0.0)
721	24	8	(33.3)
732	29	3	(10.4)
742	24	5	(20.8)

---

[1] From school permanent data provided by building principals. Based on Iowa Basic Skills Test Scores in grades four through six (May 1967 or September 1967), California Reading Test scores in grade seven (September 1967).

Table 5:

Summary of Sex Distributions

(By Grade and Class)[1]

Grade, Class	Boys	Girls	Total N
411	10	18	28
421	12	7	19
432	12	16	28
442	12	12	24
511	10	11	21
521	10	13	23
532	7	11	18
542	10	12	22
611	14	12	26
621	14	10	24
632	13	14	27
642	9	11	20
711	10	11	21
721	9	13	22
732	10	8	18
742	13	16	29
Total	175	195	370

---

[1] Based on classes after pupil elimination because of incomplete pre-test data.

Table 6

Distribution of Pupils By Sex
and Treatment Groups

	Instructional Group		
Grade	Boys	Girls	Total N
4	22	25	47
5	20	24	44
6	28	22	50
7	19	24	43
Total	89	95	184
	Non-Instructional Group		
Grade	Boys	Girls	Total N
4	24	28	52
5	17	23	40
6	22	25	47
7	23	24	47
Total	86	100	185

Table 7

Lorge Thorndike Verbal Intelligence Scores

(By Grade Level and Condition[1]

Grade Condition	N	Mean	S. D.	Range
4 Instr.	47	108.51	14.37	76-141
4 Con.	52	106.00	16.84	76-147
5 Instr.	44	105.84	13.61	80-138
5 Con.	40	108.73	15.84	74-143
6 Instr.	50	107.58	13.60	73-136
6 Con.	47	104.89	14.51	73-137
7 Instr.	43	112.16	13.27	84-135
7 Con.	47	111.62	13.69	69-138

---

[1] In grades four through six, Level 3, Form A, Verbal; in Grade Seven, Level 4, Form A, Verbal. Based on final number of pupils with complete pre-test data.

to the control condition. At the sixth-grade level, classes from Moravia and Cincinnatus were assigned as if from the same school system (that is, one class served in the experimental condition and one served in the control condition). Table 6 summarizes the distribution of pupils by sex and treatment groups. Table 7 summarizes the IQ scores by grade level and condition.

The Experimental Instructional Materials

Pupils in classes in the instructional condition studied the Productive Thinking Program developed by Covington, Crutchfield and Davies (17). This program consists of 16 lessons designed to develop the pupil's problem-solving skills and attitudes. It was originally designed to be used as a self-instructional program, and has been used entirely in that way for this study. It should be pointed out, however, that the authors of the program now feel that the materials will not be optimally effective if used on an entirely self-instructional basis.[3] Pupils in the experimental condition in this study were given one of the 16 booklets in the program on each of 16 consecutive school days in accord with the procedures followed during previous research with the program. Each booklet required about 40 minutes for completion. It should also be noted that the authors of the program have recently recommended that the materials should

---

[2] Dr. Robert Olton, personal communication, March, 1968.

be used in a less concentrated fashion in the classroom, perhaps even expanding the period of study to as long as eight weeks.

The Productive Thinking Program has been described in extensive detail in reports by Covington and Crutchfield (16) and by Covington, Crutchfield and Davies (18). Thus, for the requirements of this study, the program will be only briefly described.

The 16 booklets comprising the Productive Thinking Program present the pupil with a series of mystery of detective problems. The pupils are first introduced or the characters of the story: Jim and Lila, typical upper elementary school age youngsters, and their Uncle John, a high school teacher in the town of Elmtown. Jim and Lila have come to Elmtown to live with Uncle John while their parents go to Africa to work in the Peace Corps. The children soon see "Mr. Search," a mysterious detective who assists the local police department in solving difficult cases. They become interested in learning his identity. They soon discover that Uncle John is Mr. Search. This story line serves as the foundation for instruction in problem solving skills and attitudes. In order to teach Jim and Lila to become better thinkers, Uncle John leads them through a variety of problem situations. The pupil studying the Productive Thinking Program is assumed to identify with the characters and to become involved in their attempts to learn to be better thinkers. The principle of active involvement in the learning situation is established by asking the pupil to respond to the questions and problems as if he were in the same situation as Jim and Lila. "Feedback" is provided through extensive dialogue between Uncle John and the youngsters in the story, although emphasis is always placed on the value of independent thought and the possibility of a number of "correct" answers.

The pupils encounter 16 "thinking guideposts" which are systematically introduced, utilized by the characters, and reviewed. There are also two opportunities for "self-tests" within the programed instructional sequence. The pupils are given opportunities to produce ideas, solve problems, evaluate the characters' answers and attitudes, and to express their own attitudes.

Thus, the content of the instructional materials is independent of the subject areas ordinarily included in the school curriculum. The program attempts to develop skills, abilities and attitudes about creative problem solving through direct instruction.

The "thinking guideposts" presented in the program are:

(1) Get the facts well in mind; reflect on the problem.

(2) Decide what to work on first.

(3) Be planful in your work.

(4) Don't jump to conclusions; keep an open mind.

(5) Think of many ideas.

(6) Think of unusual and clever ideas.

(7) Pick out each main person and object in the problem.

(8) Use the "Idea Tree" (of main and particular ideas).

(9) Almost anything can remind you of ideas.

(10) Check ideas against the facts.

(11) Pay attention to puzzling facts.

(12) Try to explain puzzling facts.

(13) Try to find one idea that explains everything.

(14) Review the facts.

(15) Look at the problem in a new way.

(16) "Just suppose..." (that such and such is the case; how could it have come about?)

The authors of the program contend that its content is "atheoretical," in that it is not the product of, and is not dependent upon any theoretical formulation of problem solving. It is readily apparent, however, from the thinking guideposts, that the program's content is generally compatible with recent theoretical work on problem solving (e.g., 39, 41).

Procedures

Following the testing and data collection dealing with class selection procedures, all pupils were pre-tested on verbal creativity and arithmetic skills. Pupils in grades four, five, and six were pre-tested in October and November 1967. Pupils in grade seven were pre-tested in February 1968.

Immediately following the pre-testing period, pupils in the instructional condition classes began their work with the Productive Thinking Program. The instructional sequence, as noted above, consisted of one school period on each of sixteen consecutive school days. Control condition classes received only their ordinary classroom instruction.

Final testing of all pupils was conducted on three successive days immediately following the sixteenth day of instruction for the experimental groups. In grades four, five, and six, post-tests were given in December 1967. In grade seven, post-tests were given in March 1968. The post-tests consisted of the following measures:

1. the Torrance Tests of Creative Thinking, Verbal Form A (68);
2. the General Problem Solving Test;
3. one of the two forms of the Arithmetic Problem Solving Test (each form given to half the pupils in a class, randomly selected);
4. the Make Up Problems Test (34);
5. a slightly modified form of the Childhood Attitude Inventory for Problem Solving (14).

Measuring Instruments

Each of the measuring instruments used in this study will now be discussed in detail. Evidence for their reliability and validity will also be presented and discussed.

1. Tests of Verbal Creativity. The measures of verbal creativity used were the Torrance Verbal Tests of Creative Thinking, Form A (68). This battery consists of seven sub-tests administered in one session in a paper and pencil format. Approximately 45 minutes are required for administration. Six of these sub-tests and several combined scores were used in the present study.

The sub-tests used were:

(a.) three Ask and Guess tests (Asking Questions [AQ], Guessing Causes [GCa], and Guessing Consequences [GCo]);

(b.) Unusual Uses of cardboard boxes (UU);

(c.) Product Improvement of toy elephant (PI);

(d.) Just Suppose clouds had strings (JS).

Each sub-test was scored for one of the three components of the operational definition of creativity discussed in Chapter Two. Asking Questions and Unusual Uses were scored for flexibility. Guessing Causes and Product Improvement were scored for originality. Guessing Consequences and Just Suppose were scored for fluency. The decision to score each sub-test on only one of the three scoring dimensions was largely influenced by the practical consideration of economy in time and expense of scoring. The Ask and Guess tests were scored in the same pattern that Dacey used (24). The Unusual Uses task was scored for flexibility, in view of Torrance's emphasis that the task challenges the subject's ability to free his mind of a well established set (68, p. 12). The Just Suppose task was scored for fluency in view of the fact that Torrance considers it a "variation of the Guess Consequences task" (68, p. 13), which is scored for fluency (following 24, 57). The Product Improvement task was scored for originality. Torrance suggests that this task permits subjects "to play with ideas that they would not dare express in a more serious task," which suggested to the writers an emphasis on originality.

For the purposes of analysis, ten scores were used. First, each of the six sub-tests was used separately. Next, three combined scores (one each for fluency, flexibility, and originality) were used. Each of these was based on the arithmetic sum of the two tests scored for that dimension (following Torrance, 68). Finally, a composite total score was computed. Each of the six sub-test scores was converted to a standard score, based on a mean of 50 with a standard deviation of 10. (Standard Scores were computed separately for each of the four grade levels). The _total_ score for each pupil was the arithmetic _mean_ of the six standardized sub-test scores.

Reliability. Table 8 summarizes the test-retest reliabilities for each of these ten scores over a three week period. These data are from control classes only, since no special instructional program assumed to relate to scores on such tests was provided for these pupils.

All Fluency and Flexibility tasks were scored by two scorers, trained by the writer, using procedures recommended by Torrance (68). All originality scoring was done by one of these scorers, to reduce possible inter-scorer variation on that scoring dimension. Table 9 summarizes the mean and standard deviations of each scorer for all sub-tests, and presents the coefficients of inter-scorer reliability (Pearson product-moment correlation coefficients) based on a sample of tests scored by both scorers. Table 10 presents scorer stability coefficients for originality sub-tests based on a sample of tests re-scored by the same scorer at the end of scoring (since the same scorer scored all originality sub-tests.) These data suggest that the tests provided reliability levels sufficiently great to warrant the use of the tests in this study.

Validity. Evidence for the validity of the Torrance Tests of Creative Thinking has been presented and discussed by Torrance (68). This problem of validity has four principal aspects: content validity, concurrent validity, predictive validity, and construct validity. The evidence presented by Torrance for the validity of the tests in each of these four aspects will be presented. Data from the present study will also be presented to supplement the evidence cited by Torrance.

Torrance holds that the tests have content validity as a result of his consistent and deliberate effort "to base the test stimuli, the test tasks, instructions, and scoring procedures on the best theory and research now available." (68, p. 24). The theory and research which led to the development of the tests included analyses of the lives of indisputably eminent creative people, the nature of creative performance, and research and theory concerning the functioning of the human mind. The test tasks are, further, free of technical or subject matter content. The complexity of creativity and thepresent state of test development does not make it possible to completely specify or assess creative thinking abilities.

Table 8

Test-Retest Correlations: Torrance Tests of Creative Thinking (Control Pupils Only)

	Grade 4		Grade 5		Grade 6		Grade 7	
Subtest	N	r	N	r	N	r	N	r
AQ	52	.62**	39	.59**	45	.45**	44	.43**
GCa	52	.52**	39	.16	45	.23	44	.14
GCo	52	-.12	39	.52**	26	.35*	44	.46**
PI	52	.61**	39	.59**	45	.61**	44	.34*
UU	52	.36**	39	.31	45	.60**	44	.46**
JS	52	.56**	39	.28	26	.58**	44	.27
Flex. Sub-Ttl.[1]	52	.57**	39	.56**	45	.66**	44	.53**
Flu. Sub-Ttl.[2]	52	.33**	39	.64**	26	.48**	44	.40**
Orig. Sub-Ttl.[3]	52	.66**	39	.58**	45	.59**	44	.30*
Standardized	52							
Total Score[4]	52	.65**	39	.69**	26	.54**	44	.52**

---

[1] Asking Questions and Unusual Uses

[2] Guessing Consequences and Just Suppose

[3] Guessing Causes and Product Improvement

[4] Each of the scores in the six sub-test distributions was converted to a standard score, with $\bar{X}$=50 and S.D.=10 for each grade level. Total score was derived as the sum of the six standard scores, divided by six.

*=p<.05
**=p<.01

Table 9

Inter Scorer Reliability: Torrance

Tests of Verbal Creativity[1]

Sub-test	Scorer One $\bar{X}$	SD	Scorer Two $\bar{X}$	SD	r
Asking Questions	4.72	2.25	4.16	2.57	.73**
Guessing Consequences	3.32	2.13	3.16	2.22	.78**
Unusual Uses	6.64	3.27	6.40	2.85	.91**
Just Suppose	5.48	3.70	5.20	3.62	.97**
Flexibility Sub-Total	11.36	4.40	10.56	3.87	.88**
Fluency Sub-Total	18.72	4.96	8.36	5.01	.88**

---

[1] Based on a selected sample, N=25; Fluency and Flexibility subtests only were scored by two scorers.

*=p<.05
**=p<.01

Table 10
Coefficients of Stability of Scoring
for Originality Sub-tests[1]

Sub-test	Time 1		Time 2		r
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	
Guessing Causes	3.00	3.36	3.17	3.40	.916**
Product Improvement	5.26	5.34	5.69	7.03	.941**
Originality Sub-total	8.26	6.83	3.17	6.71	.965**

---

[1] Based on a randomly selected sample of tests, N=25. Stability coefficients for other sub-tests were also obtained from this scorer's work. These ranged from .916 to .995.

**=p<.01

Torrance recognizes this, although he does contend that, while the test tasks do not sample the entire universe of creative abilities, the Torrance tests do sample a wide range of such abilities (68, pp. 23-24).

Evidence for the concurrent validity of the tests is difficult to assess, since there are no generally accepted criteria for such validation. A frequently used criterion is peer nominations or sociometric ratings. Yamamoto found that performance on selected sub-tests from the Torrance batteries correlated positively and significantly with sociometric ratings obtained from secondary school pupils. (79, 82). Torrance also reports data from 33 elementary school classes (grades one through six). He reports that children who tended to be perceived by other children as having good ideas also tended to score higher on the Ask and Guess tests in the Torrance battery.

Another criterion with which pupils' scores have frequently been compared is teacher nominations. Torrance reports several studies (63, 65, 71, 80), in which elementary and junior high school pupils' scores were studied in relation to teachers' ratings. In these studies, pupils nominated by teachers as highest on fluency, flexibility, originality, and elaboration and those nominated as lowest on these dimensions were differentiated by the appropriate scores on the Torrance tests. Torrance and Gupta found that teachers were able to differentiate on fluency, flexibility, and originality, but not on elaboration (70). Similar results were reported by Yamamoto (80), and Nelson (48) also reported similar results for verbal test batteries, with sixth graders.

In the present study, teacher ratings were also obtained. Teachers were provided with a definition of creative behavior developed by Dacey (24), and asked to observe their pupils accordingly for a one week period. Each teacher was then provided with a class list, and was asked to rank the pupils in his class (following a procedure previously employed by the writer (73). Copies of the definition and the ranking procedure are included as Appendix J. Correlations between Standardized Total Creativity Scores (Pre-Tests) and the teachers' rankings have been computed for thirteen of the sixteen teachers whose classes participated in the study. These correlations, summarized in Table 11, ranged from -104 to .60. The mean correlation, computed using Fisher's Z-transformation procedure (33) was .37.

Torrance (68, pp. 46-7) also reports research which demonstrated a positive relation between performance on the Torrance Tests of Creative Thinking and sales productivity among department store personnel (75, 76).

Table 11

Correlations Between Teachers' Rankings of Pupil Creativity and Pupils' Standardized Total Creativity Pre-Test Scores

Class	N	r
411	28	.47*
421	19	.38
442	24	.57**
511	21	.47*
532	18	.49*
542	22	.48*
611	26	.31
621	24	.39
632	27	.29
711	21	.60**
721	22	.12
732	18	.23
742	29	-.04
Average[1]		.37

---

[1] Average r obtained using Fisher's Z-transformation Garrett, 1958

*p<.05
**p<.01

Bish found a positive relationship between scores on the verbal form of the Torrance tests and California Achievement Test scores among fourth, fifth, and sixth graders (4). The correlations ranged from .36 to .42, were all statistically significant, and increased when IQ was partialled out. Similar results were reported by Cicirelli (12), and by Torrance and Yamamoto in a study at Minnesota (68, p. 48). Bowers also reported that pupils' scores on the TTCT were positively related to educational achievement among ninth graders (6).

Only a limited amount of evidence is available concerning the predictive validity of the Torrance tests, although the author reports that a variety of long-range studies are underway or planned. In one recent project 66 seniors at the University of Minnesota High School were studied (30). These pupils were first tested in 1959, and then followed up in 1966. Erickson constructed a checklist of creative activities for the follow up measure. Based on a sample of 44 of the 66 pupils, creativity scores from 1959 were correlated with the checklist criteria. Fluency and Flexibility scores correlated positively and significantly with the checklist criterion measure (.27 and .24, respectively). The 44 subjects were then divided into two equal groups on the basis of creativity scores. Tetrachoric correlations were computed between the creativity scores and the checklist items. Originality, Elaboration, and Total Creativity scores predicted successfully participation in a number of activities (e.g., writing a poem, story, song, or play; learning a new language; changing religious affiliation; receiving a research grant; publication of scholarly paper, etc.) at acceptable statistical levels.

Torrance discusses additional evidence for the construct validity of his tests under several headings. For the present purposes, only those studies involving use of the tests with children will be reviewed. Weisberg and Springer studied fourth-grade children (78). Children scoring high and low on the Torrance tests were compared on several personality dimensions. The high-scoring children were rated significantly higher than the lows on strength of self-image, humor, ease of early recall, availability of Oedipal anxiety, and uneven ego development. These are held by Torrance to reflect "creative acceptance of oneself" and greater self-awareness. (68, p. 25). On Rorschach protocols, children ranking high on the tests showed a tendency toward unconventional responses, unreal percepts, and fanciful and imaginative treatment of the ink blots. They also gave more human movement and color responses, signs often associated with imagination and creativity in Rorschach interpretation.

Torrance studied twenty-three groups of pupils, including grades one through six (63). Three personality characteristics stood out as differentiators between highly creative children and their less creative peers. These were: first, higher scoring pupils tend to produce more wild or silly ideas; second, high scorers also tend to display a higher degree of originality in drawings; third, the higher-scoring children's productions were characterized by humor, playfulness, and relative relaxation.

Fleming and Weintraub examined the relation between Torrance test scores and rigidity among 68 gifted elementary school pupils (31). The rigidity scores correlated negatively and significantly with originality, fluency, flexibility, and composite total scores. Yamamoto reported substantial correlations (.49 and .51) between Torrance test composite scores and a measure of originality derived from imaginative stories of fifth-and sixth-grade pupils (81).

Long and Henderson used a sample of Torrance test items and found that pupils in grades two through seven who scored high on the creativity measures also tended to be better able to withhold opinions, to withstand the uncertainty of indecision, and to resist premature closure (43). Although information concerning the validity of these tests is limited in a number of ways (particularly by the absence of more extensive predictive validity data), the existing evidence warrants, in the opinion of these writers, the conclusion that the tests may be employed in this study.

2) The Lorge Thorndike Intelligence Test. The IQ pretest used in the present study was the Lorge Thorndike Intelligence Test (44). In grades four, five, and six, Level III, Form A (Verbal was used. In grade seven, Level IV, Form A (Verbal) was used. Information concerning the reliability and validity of the Lorge Thorndike tests is presented in the Technical Manual (44). Alternate Forms reliability coefficients are .896 for Level III, Verbal, and .865 for Level IV, Verbal. Odd-even split half reliability coefficients are .940 for Level III, Verbal, and .929 for Level IV, Verbal. Test-retest coefficients for Verbal tests administered to a sample of fourth-and fifth-grade pupils at a three year interval yielded a test-retest reliability coefficient of .75 (fourth) and .79 (fifth).

Evidence for the tests' validity is also presented. The authors propose that the Lorge Thorndike tests can be characterized by six statements, in general, and that these stated characteristics also apply to behavior which could be considered intelligent. These statements are:

1. The tasks deal with abstract and general concepts.
2. In most cases, the tasks require the interpretation and use of symbols.
3. In large part, it is the relationships among concepts and symbols with which examinees must deal.
4. The tasks require the examinee to be flexible in his basis for organizing concepts and symbols.
5. Experience must be used in new patterns.
6. Power in working with abstract materials is emphasized, rather than speed (44, p. 14).

Evidence for the tests' predictive validity is presented in the form of correlations of test scores with achievement test performance, using several different standard achievement tests as criteria, and in the form of correlations of test scores with school grades. In each case, the correlations were positive and of appreciable magnitude (generally greater than .60); correlations with school grades tended to be lower than with achievement measures, but were also positive and significantly different from zero.

Additional evidence for the construct validity of the tests is presented in the form of substantial correlations with other measures of intelligence. The Lorge Thorndike tests correlate positively and significantly with the California Tests of Mental Maturity, the Kuhlmann-Anderson Intelligence test, and the Otis Intelligence test. The median correlation of Lorge Thorndike scores with CTMM scores was .79 for verbal forms. For the Kuhlmann-Anderson test, the median correlation with Lorge-Thorndike scores (verbal) was .77, and .84 for Otis scores (verbal).

3. Arithmetic Pre-Test. The arithmetic pre-test used in this study was a test developed as an admissions-selection test by the members of the mathematics faculty of the Nichols School of Buffalo, New York. This test consists of 50 items, emphasizing computational skills. The test has been used for several years, with pupils in grades four, five, six, and seven. Based on the test administration data from

this study, the Kuder-Richardson Formula 20 reliability coefficient (an estimate of the average of all possible split-half reliability coefficients) was .93.

Apart from evidence for the content validity of the test, there is little evidence concerning its validity. The test was constructed by a committee of experienced mathematics teachers to represent arithmetic skills which could be reasonably expected from pupils in grades four through seven. The test has been used for several years, and the results of these test administrations have indicated an expected pattern: that scores increase as grade level increases (as found also in the present study); that the test is neither so difficult that most fourth graders score zero, nor so easy that able seventh graders readily obtain perfect scores. Its content stresses computation with all four basic arithmetic operations, and includes topics found in a wide range of arithmetic texts (whole numbers, common and decimal fractions, percentage, ratio and proportion, weights and measures, set notation, prime and composite numbers, and simpler linear equations).

Some evidence for the construct validity of the test consists of the positive correlations between test scores on this instrument and IQ scores, general problem solving test scores, and arithmetic problem solving test scores in the present study. The intercorrelations of these measures are summarized in Appendix G. The levels of difficulty and discrimination indexes for each item in the test are summarized in Appendix A, and a copy of the test is also included.

4. The General Problem Solving Test. The General Problem Solving Test consisted of eight problems presented to the pupils in paper and pencil format in a single test administration. The test was constructed by the writers for use in this study. Pupils were given forty-five minutes to work on the problems. They could attempt the problems in any order, and could return to one problem after having proceeded to another. They were instructed three times, at regular intervals during testing, not to spend too much time on any single problem. Each of the problems in the test will be described briefly; a copy of the test is included as Appendix B. The problems will be reviewed in three groups - first, those corresponding to "Type C" problems, according to the criteria proposed by Davis (25); second, those corresponding to Davis "Type O" problems; finally, a third category consisting of one remaining problem in the battery (an unsolvable problem). Numbers in parentheses refer to the placement of the problem in the test booklet.

In addition, a brief statement will be included in which each group of problems is related to the instructional content of the Productive Thinking Program.

Type C Problems. Four of the eight problems were of the kinds classified by Davis as Type C.

1. The Marble Problem (#1). This problem was taken from a study by Raaheim (53). The problem called for the examinee to transfer marbles from a cup on one table to an empty cup, on another table, six feet away. The examinee must assume a fixed position in the room (behind the table with the filled cup), and may not throw the marbles. The usual solution involves using a newspaper (provided) to form a tunnel or tube through which the marbles may be rolled to the other cup.

2. The Candle Problem (#5). This problem was first used by Duncker (27). The examinee was presented a list of available materials (a box of matches, several thumb tacks, a sheet of thin paper, a book, and a fork) and instructed to generate one or more ways of mounting three candles on a bulletin board. Several acceptable solutions are possible. The "classic" solution (obtained from a more restricted list of materials presented to older examinees) involves using the match boxes as shelves. It was also possible, from the present list, to tack the candles by their wicks, to use the fork as a shelf or as a nail-substitute, to use the paper to form an envelope, or to use the book as a shelf.

3. Water Jar Problems (#7). The problem presented in the present test was a relatively simple variant of the well-known Luchins Water Jar Problems (45), originally used to study the experimental induction of "set". The

examinee was presented two problem sets in this item. In the first part, he was presented with this situation. "You have three water cans. One holds 127 gallons; the second one holds 21 gallons; the third one holds 3 gallons. Knowing only these facts, how can you bring exactly 100 gallons of water in one trip from the water pump?"

In the second part, the examinee was presented with the following problem: "You have three cans. One holds 140 gallons; one holds 20 gallons; the third holds 10 gallons. Knowing only these facts, how can you measure out exactly 90 gallons to bring back?"

In the first part of this problem, the only feasible solution is to fill the largest jar (127 gallons), then remove *one* full volume of the intermediate jar and *two* full volumes of the smallest jar. It is immediately obvious that simple arithmetic manipulations will yield several possible solutions for the second part. This part was included for the purpose of ensuring that pupils at all grade levels would be able to attempt the problem.

4. *Nine Dots Problem (#8).* This problem has appeared in a number of discussions of problem solving and creative thinking and textbooks. The examinee is presented with nine dots, arranged so as to form three rows of three dots each when considered either horizontally or vertically. (The appearance is thus that of a square, with three dots on each of the four sides with an extra dot in the center

of the square.) His task is to connect the nine dots with four straight lines, without removing his pencil from the paper. Solution of the problem requires freeing oneself from the constraint of the apparent "squareness" - to extend a line well beyond the bottom row of dots, so as to intersect dots in two other rows with an ascending diagonal. Then the third line runs horizontally, crossing the "top" three dots. The fourth line descends diagonally through the center dot and the lower corner dot. The final solution thus resembles a triangle tipped slightly upward, with a line descending from the vertex of the upper angle bisecting the base, and continuing down to the bottom corner dot.

The writers felt that these problems would comprise a challenging opportunity for pupils to seek applications of the "Thinking Guideposts" taught in the Productive Thinking Program. Particularly, these problems would seem to demand that the pupil be flexible and able to look for new ways to reorganize familiar objects and situations. "Thinking Guideposts" in the programed instructional materials stressed such dimensions (e.g., "When you get stuck for ideas, don't give up. Try to look at the problem in a new way." "Keep an open mind; don't jump to conclusions." "Keep your eyes and your mind open to the things around you.") In several of the programed booklets, the pupil deals with situations in which illustrations of applications of these principles are presented and discussed.

Type 0 Problems. Three of the problems in the test have been classified as "Type 0" problems.

1. The Antelopes Problem (#2). This problem, similar to exercises provided most elementary school pupils as a game or "seat work", asked the examinee to construct as many English words as he could using only the letters which appear (and in no greater frequency per production) in the word "Antelopes." Proper names were excluded. Responses were scored on the basis of total number of valid productions, using The Random House Dictionary of the English Language (54) as the criterion standard.

2. Anagrams (#3). Although Davis (25) generally includes anagrams with "Type C" problem solving tasks, he indicates that the trial and error behavior can readily be externalized (a principle characteristic of "Type 0" grouping). Examinees were presented with five anagrams. Three of these could be rearranged to form more than one proper English word; there was only one possible rearrangement for each of the other two words. Score was total number of correct rearrangements produced.

3. Rule Finding Problem (#6). This problem, described by Davis (25) in his "Type 0" category, was used by Sassenrath (1963a, 1963b) in studies of anxiety and problem solving (59, 60). Directions were modified to make the problem easier for elementary school pupils to read and understand. The problem presents a series of words paired with numerals.

The examinee is told that a rule is being used to establish the numeral corresponding to each word. His task is to find the rule and then use it to write the numeral to be associated with two new words. Then, the examinee is told that the rule has been changed. He is given three sample pairs for the new rule, and then two test words. The solution rules used were: (a) N-1=X, for the first pairings, and (b) 12-N=X, for the second set (where N = number of letters in the word and X = solution numeral.) The score was total number of correct solution numerals produced (Zero to four).

These problems seemed to the writers to be relevant to the instructional content of the Productive Thinking Program in several respects. First, the programed lessons stress the need for producing many ideas. These problems give the pupil an opportunity to engage in such production. Second, the program also stresses the need for organization of ideas, and careful planning before beginning a problem as well as in the process of solving it, may lead to increased efficiency in achieving solutions. This seems also to be true in these problems: the pupil's production can readily be enhanced by a planful, systematic attack on the problem.

Unsolvable Problem. The fourth problem in the test was unsolvable. It asked the child to tell how one can "put a different number of dimes in each pocket" using exactly twenty dimes to be distributed among seven pockets. Pupils were told in the initial directions that it was possible that some problems could not be solved. For any problem believed to be unsolvable, pupils were directed to write "Can't be done" and to give a simple explanation to justify their assertion. The child could demonstrate insolvability in this problem (assuming that the solution universe consists of the set of all integers) simply by summing the first seven natural numbers, 1-7 (sum of 28; too many) and then the first seven integers, 0-6 (sum of 21; also too many). The use of any greater integer requires the multiple use of a smaller integer.

Since the Productive Thinking Program encourages the pupil to consider a wide range of alternatives, and to evaluate the possible consequences if the opposite of a given assumption is true, and to consider unusual possibilities, it was felt by the writers that an unsolvable problem would present a suitable challenge - first to consider the possibility that zero dimes would be a permissable entry for one pocket, and then to consider the possibility that no real solution is possible.

Reliability. No data are available concerning the reliability of the total test. There are several reasons for this: (1) Practical considerations made test-retest reliability an inappropriate index (e.g., no way to prevent pupils from discussing problems between administrations or continuing to work on problems they could remember); (2) No parallel forms of the test were available; (3) Since the test consists of at least two distinct sub-sets of items (five dichotomously scored, and three continuously scored), of varying item difficulty, and since no single composite score has been computed, the appropriateness of customary internal consistency coefficients was questionable. Such indexes generally assume a larger item pool, items of moderate and nearly equal difficulty, and the absence of time constraints (Gullikson, 1950).

Several authors have proposed, however, that problem solving research in general would benefit from the use of multiple criteria, involving the presentation of more than a single problem to examinees (26, 32). For this reason, it seems reasonable to assume that multiple criteria, as utilized in the present study, will yield a more stable assessment of criterion performance than possible when only a single problem is used.

Validity. Because these test items have not been used extensively with elementary school populations, evidence for the validity of the test is limited. In addition, there are no generally accepted tests of elementary school pupils' problem solving ability which could provide external criteria of validity.

There are some indications, however, supporting the content validity of the test. The problems' sources and use in other problem solving investigations have been identified in the descriptions of the problems. The problems can be fitted without difficulty to the criteria for Type O and C problems proposed by Davis (25), and many of them were included in that review. In addition, it has been pointed out that there are logical relationships between these problems and the general instructional content of the Productive Thinking Program.

Some evidence supporting the construct validity of the test can also be presented. The test items are positively correlated with other criteria of intellectual performance, including Lorge Thorndike IQ scores, the Arithmetic pretest used in this study, and verbal creativity scores. A matrix of correlation coefficients among these variables is included as Appendix G.

In addition, appreciable positive correlations were found between the items in the General Problem Solving Test and other problem solving criteria used in this study. These correlation coefficients are also presented in Appendix G.

While these data do not comprise conclusive evidence for the validity of the General Problem Solving measure, the present writers consider these data encouraging, and hold that these problems constitute useful, albeit preliminary, indexes of problem solving abilities among elementary school pupils.

5. Arithmetic Puzzles Test. This test consists of ten problems, each scored dichotomously (solved - not solved). The problems are similar to those used for classroom games or enrichment. The ten puzzles were taken from Spitzer's Practical Classroom Procedures for Enriching Arithmetic (62), in which grade levels are recommended for each problem. The grade level recommendations for the ten problems used in this battery are summarized in Table 12. Six of the ten were listed for grade four; nine of the ten for grade five; seven of the ten for grade six; and, three of the ten for grade seven. The problems were presented as a mimeographed booklet. Directions were read aloud by the test administrator. Pupils could attempt the problems in any order, and were asked to do all necessary work directly in the test booklets. Timing was ample, so that most pupils at every grade level were easily able to attempt every problem.

Discussion of the test with participating teachers following the use of the test gave no indication that any of the problems had been previously used by any participating teachers. A copy of the test is included as Appendix D.
Reliability. Since the items were dichotomously scored, and a single total score on the test was also computed for each pupil, an internal consistency index constitutes an appropriate measure of the reliability of the test. The Kuder-Richardson Formula Twenty Coefficient (38),

Table 12

Arithmetic Problem Solving

(Puzzle Form)

Placement of Problems By Grade Level[1]

Problem	Grade 3	4	5	6	7	8
1		X	X	X		
2			X	X	X	X
3	X	X	X			
4	X	X	X			
5			X	X	X	X
6			X	X		
7		X	X	X		
8	X	X	X			
9	X	X	X	X		
10				X	X	X

---

[1] Source: Spitzer (1956).

an average of all possible split-half reliability coefficients, was computed. This coefficient was .49. The low reliability coefficient, probably influenced by the generally high difficulty of the items, suggests that considerable caution should be exercised in interpreting results utilizing this test.

Validity. Very little evidence is available to provide support for the validity of the test. No predictive validity data are available. One measure of concurrent validity is the correlation between pupils' scores on this test and their scores on the Arithmetic Pre-test used in the study. These measures correlate significantly and positively ($r=.54$, $p<.01$). Evidence for the content validity consists primarily of the sources of the problems: books devoted particularly to (and generally accepted as content valid) arithmetic problems. Other evidence for construct validity consists primarily of appreciable positive correlations with Lorge Thorndike IQ scores, verbal creativity scores, and other problem solving criteria used in this study. These data are summarized in Appendix G.

6. Arithmetic Problems - Text Form. This test consisted of sixteen problems, each dichotomously solved (correct - incorrect). The pupils were presented with the test booklets and asked to attempt to solve each problem. Pupils could work on the problems in any order. Sufficient time was allowed so that most pupils at all grade levels could attempt every problem. The problems were selected

from several commercial arithmetic textbook series, from grades four through eight. Several problems were selected from texts at each grade level, so that the same test could be used at each of the four grade levels in the study. Eighth-grade texts were also included to prevent a "ceiling effect" for seventh graders of high mathematical ability. Names were changed in all problems, as were digits (where this could be achieved without changing the nature of the problem or the arithmetic required for solution). Subsequent discussion with participating teachers gave no evidence that any problem in the test had specifically been encountered and recalled by pupils. The score used for analyses was total number of problems correctly solved. A copy of this text is included in Appendix C.

Reliability. For this test, the Kuder-Richardson Formula Twenty reliability coefficient was computed. The average of all possible split-half reliability coefficients, computed by this formula, was .83. The magnitude of this coefficient offers support for the generally acceptable reliability of the test.

Validity. As was the case with both the General Problem Solving Test and the Arithmetic Puzzles Test, there is very little evidence available concerning the validity of the Arithmetic Problem Solving Test - Text Form. Support for the content validity comes from the knowledge that test items were selected from recognized arithmetic text books, appropriate for the grade levels included in this study. The general pattern of increasing mean scores as a function of grade level suggests - although tentatively - that item selection did sample an acceptable range of levels. Evidence for the concurrent validity of the test consists of an appreciable positive correlation between scores on this test and scores on the Arithmetic Pre-Test ($r=.85$, $p<.01$). Other evidence for the construct validity of the test, in the form of positive correlations with other measures of intellectual ability, including the Lorge Thorndike IQ test, the Torrance Tests of Creative Thinking, and the other problem solving measures used in this study, is also summarized in Appendix G.

7. Make Up Problems Test. The Productive Thinking Program seeks to develop a number of skills, abilities, and attitudes which are related to problem sensitivity and effective organization of data in working on a problem. The pupil is encouraged to be planful, to get the facts well in mind, to decide what to work on first, and to try to think of many ideas.

The Make Up Problems test developed by Getzels and Jackson (34) seemed to call for the application of such principles. In this test, the pupil was presented with four paragraphs, each containing several numerical statements about common activities (including building a house and filling a swimming pool, for example). For each of the paragraphs, the pupil was directed to make up as many arithmetic problems as he could. The test thus called for the child to organize the data presented, to try to find ways to use the facts presented, and to work "planfully."

Procedures for scoring the test were described in detail by Getzels and Jackson (34, pp. 205 - 208). In brief summary, a child's score on each paragraph represents the sum of the number of elements ("bits" of numerical information from the given data) and the weighted number of mathematical operations (one point for each use of addition and subtraction and two points for each use of multiplication and division). The only restriction given to pupils was that the problems written must be able to be solved without additional factual information. The pupil's total score was the sum of the points earned on each of the four paragraphs. A copy of this test is included as Appendix E.

Reliability and Validity. Getzels and Jackson (34, p. 208) report an internal consistency reliability coefficient for this test of .81, based on a sample of 45 pupils from their study. The authors estimated scorer reliability by having two judges independently score 50 problems chosen at random. The product moment correlation between these two sets of judgments was .91. In the present study, the product-moment correlation for the stability of a single trained scorer, who scored 42 randomly chosen test protocols twice independently, was .966.

Only a limited amount of evidence is available to support the validity of the test. Getzels and Jackson reported that it correlated appreciably and positively with a number of other criteria associated with creative performance and intellectual ability among pupils in their sample (34, pp. 20-25). For girls, the test correlated from .269 to .525 with several other creativity test scores, and .393 with IQ. For boys, the test correlations with other creativity criteria ranged from .175 to .420, and scores correlated .246 with IQ. Scores on the Make Up Problems test also correlated positively with verbal achievement (r=.524 for boys, .604 for girls), and with numerical achievement (r=.302 for boys and .407 for girls). In view of the restricted range of the Getzels and Jackson sample with respect to IQ, these correlations may be considered conservative estimates of the relationships among the variables considered. Considering data from the present study, appreciable, positive correlations were also found between Make Up Problems scores and other criteria (see Appendix G).

8. Childhood Attitude Inventory for Problem Solving. This paper and pencil test, developed by the staff of the Berkeley Creativity Project (14), consisted of two parts. The first part dealt with the pupil's attitudes about creativity and the nature of the problem solving process, and contained 30

statements. The pupil responded yes or no to indicate agreement or disagreement with each statement. The second part, also 30 "yes-no" statements, dealt with the pupil's self-confidence about engaging in creative problem solving activities. In the present study, a preliminary form of this attitude measure was used which was slightly different from the later ("revised") measure. Part I consisted of 30 items, 28 of which were identical with the revised version. Part II consisted of only 22 items, all of which also appear in the revised version. The child's score on each part was the total number of responses which express favorable attitudes concerning creative problem solving. In addition to the pupil's scores on both parts of the measure, the present study also used a Total score (the sum of a pupil's scores on Parts I and II.) A copy of this test is included in Appendix F.

Validity and Reliability. Evidence concerning the validity and reliability of the test has been presented by Covington (1967). He reported data from 325 fifth-and sixth-grade pupils. Test-retest reliability coefficients, over a five-week interval, averaged .69 for Part I and .65 for Part II. Covington also reported Kuder-Richardson Formula 20 Reliability coefficients for the preliminary form of the test (i. e., for the form used in this study). These were .93 for Part I and .86 for Part II. Covington also presented data to indicate that test scores were relatively unaffected by age, and that test scores (for both parts) correlated significantly and positively with an intelligence measure (CTMM), and were negatively correlated with anxiety indexes. These data provide preliminary evidence for the construct validity of the test.

The reliability coefficients for all instruments used in this study are presented in summary, in Table 13.

Treatment of the Data

The treatment of the data will be discussed in three sections. First, analyses in which instructional and control pupils were compared with respect to each of the several creative problem solving criteria discussed above will be discussed. Then, the procedures for the "internal analyses" in which pupils' responses from the programed instructional materials were analyzed will be described. In the third section, comparisons of creative problem solving scores for pupils high and low on each program factor (and with a sub-sample of control pupils) will be discussed.

Principal Instructional-Control Comparisons. In these analyses, pupils' scores on each of the creative problem solving criteria were compared. Scores of pupils in the instructional group were compared with scores of pupils in the control group. All analyses were conducted separately for each grade level.

1. Verbal Creativity. At each grade level twenty scores were obtained for each pupil (ten scores on both

Table 13:

Summary of Reliability Coefficients
for Instruments Used in this Study

Test, Source of data	Coefficient	r
Lorge Thorndike IQ (Published Manual)	Alt. forms-level III	.896
	Alt. forms-level IV	.865
	Odd-Even, Level III	.940
	Odd-Even, Level IV	.929
Torrance Tests of Creative Thinking (Torrance, 1966)	Test-Retest	.61-.93[1]
Arithmetic Pre-test (Present Study)	Kuder-Richardson #20	.93
Arithmetic Puzzles (Present Study)	Kuder-Richardson #20	.49
Arithmetic Text Problems (Present Study)	Kuder-Richardson #20	.83
Make Up Problems (Getzels and Jackson, 1962)	Internal Consistency	.81
Attitude Inventory (Covington, 1967)	K-R #20 Part I	.93
	K-R #20 Part II	.86
	Test-Retest, Part I	.69
	Test-Retest, Part II	.65

---

[1] Range of coefficients for verbal sub-tests, grades 4 - 6.

pre- and post-tests). The ten pre-and post-test scores were: Asking Questions, Guessing Causes, Guessing Consequences, Product Improvement, Unusual Uses, and Just Suppose (all sub-test scores); Fluency, Flexibility, and Originality (all sub-totals based on two sub-tests each); and Standardized Total Score (a composite derived from all six subtest scores).

For each of these ten sets of scores, two analyses of covariance were conducted (29, 35). In both analyses, post-test scores were used as the variate (dependent variable). In the first analysis, the pre-test score for the appropriate criterion was used as a covariate (or concomitant variable), since pre- and post-test scores were significantly correlated at each grade level. In the second analysis, both pre-test scores and IQ scores were used as concomitant variables. The use of IQ as a covariate seemed justified by previous research which has indicated that, for groups which are heterogeneous with respect to IQ, a moderate positive correlation with verbal creativity would be expected (41). In addition, the dependent variable scores were significantly and positively correlated with IQ scores at all four grade levels in the present study.

Thus, at each grade level, and for each of the ten verbal creativity post-test scores, instructional and control group means were compared, first controlling statistically for the effects of pre-test scores, then controlling statistically for the effects of both pre-test and IQ scores.

2. General Problem Solving. Five analyses were conducted to compare instructional and control pupils' scores on the General Problem Solving Test. First, the proportions of pupils in each treatment group correctly solving each of the Type C problems were compared by grade level, using a two-by-two Chi-square test (two treatments by solving-failing to solve; Siegel, 1956).

The computational formula for χ^2 in the case of a two-by-two contingency table with 1 df., is:

$$\chi^2 = \frac{N(|AD-BC|-N/2)^2}{(A+B)(C+D)(A+C)(B+D)}$$

(where N is the number of cases and A, B, C, D, are the observed frequencies in each of the four cells of the two-by-two contingency table.)

Next, the total number of solutions for Type C problems (which takes into account the production of more than one solution per problem as well as number of problems solved) was obtained for all pupils. The Type C Total Scores of pupils in the instructional and control groups were compared using one-way analysis of covariance. Since in the present study IQ was significantly and positively correlated with Type C Total Score (product moment correlations ranging from .33 to .57, by grade level, with $p<.01$ and on the basis of the logical relation of IQ to problem solving, IQ was used as a concomitant variable (Covariate) in these analyses.

Third, the proportions of pupils in the instructional and control groups giving acceptable responses on the unsolvable problem were compared, using a Chi-square test for two-by-two contingency tables (as for the separate treatment of Type C problems, above.)

Fourth, the scores of instructional and control group pupils on each of the Type O problems were compared, using one-way analysis of covariance.

Since IQ was considered to relate logically to pupils' performance on Type O Problems, and since Type O problem scores were significantly ($p<.01$) and positively correlated with IQ at each grade level,[3] IQ scores were used as the concomitant variable of covariate.

Finally, the scores on each Type O problem were standardized by grade level to $\bar{X}=50$, S. D.=10. A standardized Total Type O Score was obtained by summing pupils' standardized scores on each of the three Type O problems and dividing by three. Then, Standardized Total Type O scores for instructional and control group pupils were compared, using one-way analysis of covariance (29, 35). In these analyses as in the previous analyses, IQ scores were used as concomitant observations or covariates.

---

[3] The only exceptions were the Rule Finding Problem in grade five ($r=.20$) and the Antelopes problem in grade six ($r=.15$).

3. Arithmetic Problem Solving Tests. The analyses of scores on both forms of the Arithmetic Problem Solving tests utilized smaller samples of instructional and control pupils, since each pupil received either the Puzzle Form or the Text Form, but not both. These tests were distributed so that half the pupils in each class were randomly assigned to one Form of the test or the other.

For the Puzzle Form, two sets of analyses were conducted at each grade level. First, the proportion of pupils in each treatment group correctly solving each problem were compared, using a two-by-two Chi-Square test (Siegel, 1956). Second, the total number of correct solutions was computed for each pupil. Since, at each grade level, these total scores were significantly and positively correlated with both Arithmetic Pre-Test scores (range of product moment correlations by grade form .39 to .55, all $p<.01$) and IQ scores ($r=.26$, $p<.05$ in grade four; range of correlation coefficients from .40 to .69 all $p<.01$, in grades 5-7), one-way analysis of covariance was used to compare pupils' scores on the Puzzle Form of the test. Thus, APS-Puzzle scores of pupils in the instructional and control groups were compared, controlling statistically for the effects of IQ and Arithmetic Pre-Test.

For the Text Form, the total scores of pupils in both treatment groups were compared, at each grade level, also using one-way analysis of covariance. Since Text Form scores were significantly and positively correlated with both Arithmetic Pre-Test scores (range of correlation coefficients by grade level, .69 - .85, all p<.01) and IQ scores (range of correlation coefficients, by grade level, .57 - .80, all p<.01), both IQ and Arithmetic Pre-Test scores were used as concomitant observations. Thus, the scores of pupils in both treatment groups were compared, with IQ and Arithmetic Pre-Test scores statistically controlled.

4. Make Up Problems. At each grade level, scores of instructional and control pupils on the Make Up Problems test were compared using one-way analysis of variance.

5. Attitude Inventory. On the Childhood Attitude Inventory for Problem Solving, three analyses were conducted at each grade level. One-way analysis of variance was used to compare instructional and control pupils' mean scores on: (1.) Part I, dealing with attitudes about creative problem solving; (2.) Part II, dealing with self-confidence about engaging in creative problem solving; and (3.) Total Score, the sum of scores on Parts I and II.

Internal Analyses. In previous research with the Productive Thinking Program, these instructional materials have been treated as if they constituted a single dimension or unitary experimental "manipulation."

In the sense that the program provides an organized, logically developed, instructional sequence such a procedure is certainly warranted, and such comparisons have been utilized in this study.

It seemed important, however, to investigate the instructional content of the Productive Thinking Program more closely. As a result, an attempt has been made to identify certain components or factors, which together comprise the total instructional sequence. Since this has not been attempted in any previous research with these materials, these efforts are viewed by the writers as exploratory.

1. Categorization of pupils' program responses. The first decision in designing this exploratory investigation was that the pupil responses required by the program constituted an appropriate domain in which to work. Thus an attempt was made to categorize the instructional content of the written responses which pupils are required to make. This task was undertaken by the writers with the assistance of a graduate student concerned with the psychological study of human problem solving and acquainted with the instructional materials.[4] It was decided that the pupils' program responses could be summarized using five categories. These were:

---

[4] The writers acknowledge with thanks the contribution of Mr. Frederick T. Bail.

1. Memory - Organization. This category included items which asked the pupil to recall previously given facts, solutions, or thinking guideposts; items which called the pupils' attention to puzzling facts, or reviewed the facts presented in a given problem; items which required the pupil to organize given ideas, or to collect given ideas into logical categories.

2. Production. These items required the pupil to produce a response, which may be a statement of solution, an idea about the problem situation, or a generalization from available facts. The production may be convergent (where the item's constraints lead the pupil to a single, well-defined response), or divergent (where the child is asked to produce a number of responses). The emphasis was on generating responses independently, rather than merely reproducing given data.

3. Redirection. These items sought to evoke from the pupil a response which involved an entirely new or different view of the problem. Rather than asking the pupil to produce a new idea, these items required large-scale reorganization of one's thinking about the entire problem. Responses stressing openness, tolerance for ambiguity, and recombination of thoughts were included. [Several of the thinking guideposts presented in the program illustrate redirection: #15, Look at the problem in a new way; #16,

Just suppose that what you have thought to be true up to now is really false; what new ideas does this suggest?]

4. Judgment and Evaluation. This factor included items which required the pupil to give his solution idea for a problem, or to evaluate the truth or likelihood of an idea which was presented, or to select items from a given list which were inconsistent with newly presented facts.

5. Attitude. In these items the pupils were to respond in such a way as to express an affective response to a situation. He may be asked to make an affective response toward a problem (too easy or hard? fun to do? etc.), or a about the problem situation (should one give up now?, etc.), or about himself, his own abilities, or the flaws and shortcomings of the characters in the lessons (e.g., 'Jim is nice but not too smart.')

After these categories had been established, each response which the pupil was required to make in the sixteen lesson sequence was categorized. From a total of one hundred seventy-five responses, two trained raters were in agreement on the categorization of eighty-two percent of the responses. For all others (the remaining eighteen percent), the raters discussed the responses and arrived at a mutual decision about categorization. There were no responses which the raters were unable to categorize. Economy of scoring dictated that each response be assigned only to one category. A summary of the responses and assigned categories appears in Appendix H.

Following the assignment of pupil program responses to categories, the raters scored the answer booklets for 20 pupils from each of the four grade levels. These pupils were randomly drawn from the total number of pupils available at each grade level. Item responses for which pupils made more than one response were credited with one point *per response*, except in the Judgement-Evaluation factor, where a pupil could receive no more than one point per item. For each pupil, a *score* on each factor was obtained: the total number of points he had received for all items scored on that factor. Thus, each pupil received *five* scores.

Each of the two raters scored a randomly-drawn sample of protocols that had previously been scored by the other rater, in order to estimate inter-scorer reliability. On each of the eight protocols so rescored, there was greater than 80 percent agreement between raters. Inspection of the responses that were not in agreement revealed no systematic biases. Most response disagreements were no greater than one point in magnitude, and neither rater scored consistently higher or lower than the other.

2. *Analyses of the Pupil Response Data.* These scores were then used to assess the relative effectiveness of the five

response factors (considered to be components of the instructional content of the total Program) with respect to several creative problem solving criteria. First, the question was raised of the relationship of each factor to the pupils' General Problem Solving and Verbal Creativity scores. (Would there be a significantly greater [correlational] relationship between pupils' program response scores and the problem solving or creativity criteria for certain factors than for others? In addition, would pupils who were High on one response factor, rather than another, perform significantly better on the problem solving and creativity criteria than either those who were Low on the factor or than control pupils, when differences in IQ were statistically controlled?)

At each grade level, then, the correlations between pupils' scores on each of the five response-factors were correlated with each of three creative problem solving criteria: Type O Problem Solving Total Score; Type C Problem Solving Total Score; and Verbal Creativity Post-test Total Score. The differences among these correlations were tested for significance, at each grade level and for each of the three dependent variable categories (33).

Next, the pupils highest and lowest on each response factor were identified. The upper seven pupils at each grade level, on each response factor, constituted the High group and the lowest seven pupils, the Low group. Ties among pupils at high and low division points were resolved by identifying groups which would be nearest in size to seven (i.e., a group of six would be identified, if necessary, in preference to a group of nine, to maintain discrimination between high and low groups). A randomly selected sample of ten control pupils at each grade level was also identified. Then, separately for each grade level, and for each response factor, the High, Low, and Control pupils were compared on each of the three dependent variables (Type O Problem Total; Type C Problem Total; and Post-Creativity Total), using one-way analysis of covariance. IQ scores were used as a concomitant observation, to control for differences among groups. Thus, for each analysis, there were three treatment groups. If significant F-ratios indicated that two or more of the three means differed significantly, individual post-hoc comparisons of means would be conducted. Since there were five response factors and three dependent variables, fifteen analyses were conducted at each of the four grade levels.

Relation to the Specific Statement of the Problem.

In Chapter One, seven specific questions were posed as the specific issues to which the study was addressed. These questions will now be restated in relation to the procedures outlined in this Chapter.

1. What abilities are developed in the program?

2. If some abilities are more effectively developed than others by the instructional materials, to what extent do the abilities taught bear differentially on pupils' problem solving performance?

These questions concern the internal analyses. Their answers have been suggested by the categorization of pupil responses into five general factors. The results of the exploratory analyses conducted using these categories will be presented in detail in Chapter Three and discussed in Chapter Four.

tested.

3. At what grade level(s), if at all, may the instructional materials be recommended for classroom use?

4. To what extent, if at all, must the instructional materials be revised, modified, or supplemented by other activities for optimal effectiveness in the classroom?

These questions will be discussed in detail in Chapter Four. Comparisons of the results of the study at each of the four grade levels which participated, and detailed examination of the procedures followed, will be discussed.

5. To what extent, if at all, does the instructional treatment (i. e., studying the Productive Thinking Program) facilitate performance on tests of verbal creativity, when pre-test scores and IQ are statistically controlled?

In discussing the treatment of the data (above), the procedure for investigating this aspect of the problem has been described. The study has attempted to answer this question by comparing instructional and control group pupils' scores on ten verbal creativity measures, using one-way analysis of covariance, controlling for the effects of pre-test, and for the effects of pre-test and IQ. Thus, the question has been addressed as the test of the following hypotheses (stated in null form) for each set of ten comparisons:

A. There is no significant difference between instructional and control group means after adjustment to control for pre-test scores;

B. There is no significant difference between instructional and control group means after adjustment to control for pre-test scores and IQ.

6. To what extent, if at all, does what is learned or developed in the instructional treatment, lead to positive transfer to general problem solving situations?

This aspect of the problem has been examined through the several comparisons of instructional and control group, pupils' scores on the General Problem Solving Test, the Make Up Problems Test, and the Pupil Attitude Inventory. Specific hypotheses tested (in null form) which were derived from this question were:

A. There is no significant difference between treatment group means, with respect to total number of solutions, when IQ is statistically controlled. (Such a null hypothesis was tested for Type 0 problems, Type 0 Standardized Total Score, and for Type C Total Score, at each grade level, using one-way analysis of covariance.)

B. There is no significant difference between treatment groups, with respect to the proportion of pupils correctly solving a specified problem. (Such a null hypothesis was tested, using a 2X2 Chi-Square test, for each Type C Problem, and for the insolvable problem.)

C. There is no significant difference between treatment group means, with respect to scores on the Make Up Problems Test. (This null hypothesis was tested, at each grade level, using one-way analysis of variance.)

D. There is no significant difference between treatment group means, with respect to scores on the Childhood Attitude Inventory for Problem Solving. (Such a null hypothesis was tested for Part I, Part II, and Total Score, using one-way analysis of variance.)

7. To what extent, if at all, does the learning form instructional treatment lead to positive transfer to arithmetic problem solving?

This aspect of the problem has been examined through several comparisons of instructional and control group pupils' scores on the Arithmetic Problem Solving Tests (Puzzle Form and Text Form). Specific hypotheses (in null form) which were derived from this question were:

A. There is no significant difference between treatment group means, after adjustment to control for the effects of IQ and Arithmetic Pre-Test, with respect to total scores on the Arithmetic Problem Solving Test (Text Form). (This null hypothesis was tested using one-way analysis of covariance.)

B. There is no significant difference between treatment group means, after adjustment to control for the effects of IQ and Arithmetic Pre-Test, with respect to total scores on the Arithmetic Problems Solving Test (Puzzle Form.) (This null hypothesis was tested using one-way analysis of covariance.)

C. There is no significant difference between treatment groups with respect to the proportion of pupils solving a specified problem on the Arithmetic Problem Solving Test, Puzzle Form. (Such a null hypothesis was tested, using a 2X2 Chi-square test, for each of the ten problems in the test.)

Table 14 summarizes the analyses which were conducted in the study.

Chapter Summary

In this chapter, the setting and design of the study were discussed. The measuring instruments used in the study were described in detail, and evidence bearing on their

Table 14:

Summary of Statistical Analyses Conducted
At Each Grade Level

Pre-test Variables	Analysis	Comparisons[1]
Lorge Thorndike IQ	1-way ANOVA	I-C
Arithmetic Skills	1-way ANOVA	I-C
Verbal Creativity		
6 sub-tests	1-way ANOVA	I-C
3 sub-totals	1-way ANOVA	I-C
1 total	1-way ANOVA	I-C
Post-test Variables		
Verbal Creativity		
6 sub-tests	1-way Covariance	I-C
3 sub-totals	1-way Covariance	I-C
1 total	1-way Covariance	I-C
General Problem Solving		
Each Type O Problem	1-way Covariance	I-C
Each Type C Problem	Chi-square	I-C
Unsolvable Problem	Chi-square	I-C
Type O Total Score	1-way Covariance	I-C
Type C Total Score	1-way Covariance	I-C
Arithmetic Problem Solving		
Puzzle-each problem	Chi-square	I-C
Puzzle-total score	1-way Covariance	I-C
Text-total score	1-way Covariance	I-C
Make Up Problems	1-way ANOVA	I-C
Attitude Inventory	1-way ANOVA	I-C
Internal Analyses		
A. Correlational		
5 Response Factors, 3 criterion scores	t-tests of differences between Z coefficients.	(I only)
B. By Response Factors		
Type O Problems	1-way Covariance	I(high-low)-C
Type C Problems	1-way Covariance	I(high-low)-C
Creativity Total	1-way Covariance	I(high-low)-C

---

[1] I=Instructional, C=Control

reliability and validity was summarized. The instructional materials used in the study were described and related to a theoretical model of problem solving and to the measuring instruments used in the study. The procedures for statistical analyses of pre-test and post-test scores of instructional and control pupils were described in detail. The procedures followed in the development of the "internal analyses", derived from categorization of written responses during the programed instructional sequence were described, and the statistical analyses of these data were described.

CHAPTER III

RESULTS OF THE STUDY

In this chapter, the results of the study are presented in detail. These have been grouped into three general categories: (a) Pre-test comparisons; (b) Principal Instructional-Control Comparisons; and, (c) Internal Analyses. Each of these categories constitutes a principal section of the chapter. Within each principal section, results are presented for each of the four grade levels (four through seven) for which data were collected.

Pre-test Comparisons

At each of the four grade levels, the instructional and control groups were compared, using one-way analysis of variance, on each of eleven pre-test variables. These variables were: Lorge-Thorndike IQ; Arithmetic Pre-test; each of the six verbal creativity sub-tests; and, each of three verbal creativity sub-total scores. The results of these analyses follow.

Table 15 presents comparisons of pre-test scores for instructional and control groups in grade four. On only one comparison, the Originality sub-total, did the groups differ significantly. In this comparison, the instructional group mean was 5.89 and the control group mean was 4.38.

Table 15
Analysis of Variance: Pre-test variables
(Grade Four)

Variable	Means Instructional[1]	 Control[2]	F[3]	P
IQ	108.51	106.00	<1	n.s.
Arithmetic	8.43	8.96	<1	n.s.
Asking Questions	3.43	3.52	<1	n.s.
Guessing Causes	2.02	1.52	1.65	n.s.
Guessing Consequences	3.87	3.94	<1	n.s.
Product Improvement	3.87	2.87	3.18	n.s.
Unusual Uses	5.74	5.79	<1	n.s.
Just Suppose	3.21	3.35	<1	n.s.
Fluency Sub-total	7.09	7.29	<1	n.s.
Flexibility Sub-total	9.17	9.31	<1	n.s.
Originality Sub-total	5.89	4.38	4.25	<.05

---

[1] N=47

[2] N=52

[3] df= 1,97

Table 16

Analysis of Variance: Pre-test Variables

(Grade Five)

Variable	Means Instructional[1]	 Control[2]	F[3]	P
IQ	105.84	108.73	<1	n.s.
Arithmetic	14.34	14.88	<1	n.s.
Asking Questions	4.98	4.78	<1	n.s.
Guessing Causes	2.14	2.30	<1	n.s.
Guessing Consequences	4.30	3.08	2.91	n.s.
Product Improvement	4.41	6.18	3.63	n.s.
Unusual Uses	6.64	6.55	<1	n.s.
Just Suppose	4.27	4.55	<1	n.s.
Fluency Sub-total	8.55	7.63	<1	n.s.
Flexibility Sub-total	11.61	11.33	<1	n.s.
Originality Sub-total	6.55	8.48	2.30	n.s.

---

[1] N=44

[2] N=40

[3] df= 1,82

Table 17

Analysis of Variance: Pre-test Variables

(Grade Six)

Variable	Means			
	Instructional[1]	Control[2]	F[3]	P
IQ	107.58	104.89	<1	n.s.
Arithmetic	25.62	18.13	25.91	<.01
Asking Questions	3.96	4.26	<1	n.s.
Guessing Causes	1.68	3.13	7.854	<.01
Guessing Consequences*	4.72	4.52	<1	n.s.
Product Improvement	4.60	5.21	<1	n.s.
Unusual Uses	7.14	8.81	4.586	<.05
Just Suppose*	4.80	3.37	5.995	<.05
Fluency Sub-total*	9.52	7.89	2.185	n.s.
Flexibility Sub-total	11.10	13.06	3.778	n.s.
Originality Sub-total	6.28	8.34	3.224	n.s.

---

[1] N=50

[2] N=47 (see*)

[3] df= 1,95

*Data based on smaller sample in control group (N=27), because of an omission in one class during pre-test administration; df= 1.75.

Table 18

Analysis of Variance: Pre-test Variables

(Grade Seven)

Variable	Means Instructional[1]	 Control[2]	F[3]	P
IQ	112.16	111.62	<1	n.s.
Arithmetic	27.21	28.60	<1	n.s.
Asking Questions	5.49	5.15	<1	n.s.
Guessing Causes	3.63	3.32	<1	n.s.
Guessing Consequences	6.28	6.81	<1	n.s.
Product Improvement	7.33	6.79	<1	n.s.
Unusual Uses	8.53	8.19	<1	n.s.
Just Suppose	6.58	5.96	<1	n.s.
Fluency Sub-total	12.86	12.74	<1	n.s.
Flexibility Sub-total	14.02	13.38	<1	n.s.
Originality Sub-total	10.95	10.11	<1	n.s.

---

[1] N=43

[2] N=47

[3] df= 1,88

The F-ratio of 4.25, with 1 and 97 degrees of freedom, indicated that the difference between means was reliably different from zero (p<.05).

At the fifth-grade level, summarized in Table 16, there were no significant differences between instructional and control group means on any of the pre-test variables.

At the sixth-grade level the instructional and control groups differed significantly on four comparisons. These data are summarized in Table 17. The instructional group mean was significantly greater than the control group mean for the Arithmetic Pre-test (25.62 vs. 18.13, F=25.91, p<.01 with 1,95 df) and for the Just Suppose test (4.80 vs. 3.37, F=5.995, p<.05 with 1,75 df). The control group mean was significantly greater than the instructional group mean on the Guessing Causes test (3.13 vs. 1.68, F=7.854, p<.01 with 1,95 df) and on the Unusual Uses test (8.81 vs. 7.14, F=4.586, p<.05 with 1,95 df).

Comparisons of the seventh-grade instructional and control group means are summarized in Table 18. There were no significant differences between treatment groups on any comparisons.

Instructional-Control Comparisons

The instructional and control groups were compared on a number of post-test variables. These included: verbal creativity (six sub-tests, three sub-totals, and standardized total score), General Problem Solving (by problem, Type C total, and Type 0 total), Arithmetic Problem Solving (Text form total score, Puzzle problems, and Puzzle form total), Make Up Problems, and the Childhood Attitude Inventory for Problem Solving. Each of these comparisons of post-test measures will be reported separately. Within the presentation of results for each measure, results for each grade level will also be reported separately.

Verbal Creativity. For each of the ten verbal creativity scores at each grade level, two analyses were conducted. The first was a one-way analysis of covariance, comparing instructional and control group means on the creativity post-test scores, using pre-test creativity scores as a covariate. These comparisons are summarized in Tables 19a-19d. The second analysis conducted for each creativity score compared instructional and control group means, using both creativity pre-test scores and IQ as covariates. These results are summarized by grade levels in Tables 20a-20d.

At the fourth-grade level, there were no significant differences between the instructional and the control groups, on any of the ten verbal creativity indexes, when only the effects of creativity pre-test were statistically controlled (Table 19a). When both creativity pre-test and IQ were statistically controlled (Table 20a), the instructional group mean was significantly greater than the control group mean only for the Originality Sub-Total Index (6.05 vs. 4.53, $F=4.581$, $p<.05$ with 1,95 df). Of the 19 comparisons in

Table 19a:
Analysis of Covariance Comparison of Instructional and Control Creativity Post-tests, Creativity Pre-test Covaried
(Grade Four)[1]

Sub-test	Post-test Means		Adjusted Means		F	P
	Instructional	Control	Instructional	Control		
Asking Questions	3.79	4.12	3.81	4.09	<1	n.s.
Guessing Causes	2.15	1.37	2.05	1.46	2.660	n.s.
Guessing Consequences	2.43	2.69	2.43	2.69	<1	n.s.
Product Improvement	4.43	2.69	4.05	3.03	2.869	n.s.
Unusual Uses	6.11	6.00	6.12	5.99	<1	n.s.
Just Suppose	3.72	3.90	3.74	3.88	<1	n.s.
Fluency Sub-total	6.26	6.58	6.29	6.55	<1	n.s.
Flexibility Sub-total	9.89	10.12	9.93	10.08	<1	n.s.
Originality Sub-total	6.57	4.06	6.04	4.54	3.927	n.s.
Standardized Total	50.49	49.65	50.21	49.91	<1	n.s.

---

[1] Instructional group, N=47; Control group, N=52; df= 1,96.

Table 19b:

Analysis of Covariance Comparison of
Instructional and Control Creativity Post-tests
Creativity Pre-test Covaried
(Grade Five)[1]

Sub-test	Post-test Means Instructional	Post-test Means Control	Adjusted Means Instructional	Adjusted Means Control	F	P
Asking Questions	5.29	4.85	5.18	4.96	<1	n.s.
Guessing Causes	2.24	1.62	2.25	1.60	1.272	n.s.
Guessing Consequences	3.69	3.95	3.44	4.22	<1	n.s.
Product Improvement	5.02	5.31	5.57	4.72	1.160	n.s.
Unusual Uses	7.33	7.44	7.37	7.40	<1	n.s.
Just Suppose	4.45	4.21	4.54	4.11	<1	n.s.
Fluency Sub-total	8.14	8.15	7.93	8.39	<1	n.s.
Flexibility Sub-total	12.62	12.28	12.55	12.36	<1	n.s.
Originality Sub-total	7.26	6.90	7.80	6.32	2.016	n.s.
Standardized Total	50.31	49.69	50.43	49.56	<1	n.s.

---

[1] Instructional group, N=42; Control group, N=39; df= 1,78.

Table 19c:
Analysis of Covariance Comparison of
Instructional and Control Creativity Post-tests
Creativity Pre-test Covaried
(Grade Six)[1]

Sub-test	Post-test Means		Adjusted Means		F	P
	Instructional	Control	Instructional	Control		
Asking Questions	5.72	6.07	5.82	5.97	<1	n.s.
Guessing Causes	3.32	3.53	3.60	3.24	<1	n.s.
Guessing Consequences*	4.49	6.15	4.48	6.17	4.997	<.05
Product Improvement	5.57	6.24	5.76	6.05	<1	n.s.
Unusual Uses	8.00	10.04	8.40	9.63	3.052	n.s.
Just Suppose*	5.72	4.42	5.29	5.20	<1	n.s.
Fluency Sub-total*	9.81	10.58	9.60	10.95	1.122	n.s.
Flexibility Sub-total	13.62	16.11	14.17	15.53	2.582	n.s.
Originality Sub-total	8.89	9.78	9.52	9.12	<1	n.s.
Standardized Total*	49.06	50.31	49.17	50.12	<1	n.s.

---

[1] Instructional group, N=47; Control group, N=45; df= 1,89

*Based on Instructional group N=47, Control group N=26, df= 1,70.

Table 19d

Analysis of Covariance Comparison of

Instructional and Control Creativity Post-tests

Creativity Pre-test Covaried

(Grade Seven)[1]

Sub-test	Post-test Means		Adjusted Means		F	P
	Instructional	Control	Instructional	Control		
Asking Questions	7.05	7.32	7.01	7.36	<1	n.s.
Guessing Causes	7.24	5.11	7.20	5.15	5.787	<.05
Guessing Consequences	8.22	8.05	8.43	7.82	<1	n.s.
Product Improvement	9.34	8.39	9.08	8.63	<1	n.s.
Unusual Uses	10.10	10.16	10.03	10.23	<1	n.s.
Just Suppose	9.10	8.80	8.87	9.01	<1	n.s.
Fluency Sub-total	17.32	16.84	17.30	16.85	<1	n.s.
Flexibility Sub-total	17.15	17.48	17.01	17.61	<1	n.s.
Originality Sub-total	16.34	13.50	15.98	13.84	1.775	n.s.
Standardized Total	50.40	49.52	50.29	49.63	<1	n.s.

---

[1] Instructional group, N=41; Control group, N=44; df= 1,82

Table 20a:
Analysis of Covariance Comparison of Instructional and Control Creativity Post-tests, Creativity Pre-Test and IQ Covaried
(Grade Four)[1]

	Post-test means		Adjusted Means			
Sub-test	Instructional	Control	Instructional	Control	F	P
Asking Questions	3.79	4.12	3.77	4.13	1.571	n.s.
Guessing Causes	2.15	1.37	2.02	1.48	2.359	n.s.
Guessing Consequences	2.43	2.69	2.36	2.76	1.065	n.s.
Unusual Uses	6.11	6.00	6.04	6.06	<1	n.s.
Product Improvment	4.43	2.69	4.04	3.04	2.986	n.s.
Just Suppose	3.72	3.90	3.68	3.94	<1	n.s.
Fluency Sub-total	6.26	6.58	6.15	6.67	<1	n.s.
Flexibility Sub-total	9.89	10.12	9.83	10.17	<1	n.s.
Originality Sub-total	6.57	4.06	6.05	4.53	4.581	<.05
Standardized Total	50.49	49.65	50.09	50.01	<1	n.s.

---

[1] Instructional group, N=47; Control group, N=52; df= 1,95

Table 20b:

Analysis of Covariance Comparison of
Instructional and Control Creativity Post-tests,
Creativity Pre-test and IQ Covaried
(Grade Five)[1]

Sub-test	Post-test Means		Adjusted Means		F	P
	Instructional	Control	Instructional	Control		
Asking Questions	5.29	4.85	5.27	4.87	<1	n.s.
Guessing Causes	2.24	1.62	2.36	1.48	2.465	n.s.
Guessing Consequences	3.69	3.95	3.68	3.96	<1	n.s.
Unusual Uses	7.33	7.44	7.46	7.30	<1	n.s.
Product Improvement	5.02	5.31	5.59	4.69	1.265	n.s.
Just Suppose	4.45	4.21	4.64	4.00	1.193	n.s.
Fluency Sub-total	8.14	8.15	8.24	8.05	<1	n.s.
Flexibility Sub-total	12.62	12.28	12.72	12.17	<1	n.s.
Originality Sub-total	7.26	6.90	7.88	6.23	2.587	n.s.
Standardized Total	50.31	49.69	50.66	49.32	1.587	n.s.

---

[1] Instructional group, N=42; Control group, N=39; df= 1,77.

Table 20c:
Analysis of Covariance Comparison of
Instructional and Control Creativity Post-tests,
Creativity Pre-test and IQ Covaried
(Grade Six)[1]

Sub-test	Post-test Means		Adjusted Means		F	P
	Instructional	Control	Instructional	Control		
Asking Questions	5.72	6.07	5.80	5.98	<1	n.s.
Guessing Causes	3.32	3.53	3.51	3.34	<1	n.s.
Guessing Consequences*	4.49	6.15	4.44	6.24	6.509	<.05
Unusual Uses	8.00	10.04	8.33	9.70	3.667	n.s.
Product Improvement	5.57	6.24	5.59	6.23	<1	n.s.
Just Suppose*	5.72	4.42	5.29	5.20	<1	n.s.
Fluency Sub-total*	9.81	10.58	9.60	10.95	1.193	n.s.
Flexibility Sub-total	13.62	16.11	14.10	15.61	3.072	n.s.
Originality Sub-total*	8.89	9.78	9.25	9.41	<1	n.s.
Standardized Total	49.06	50.31	49.12	50.20	<1	n.s.

---

[1] Instructional group, N=47; Control group, N=45; df= 1,89.

*Based on Instructional group N=47, Control group N=26, df= 1,69.

Table 20d:
Analysis of Covariance Comparison of Instructional and Control Creativity Post-tests, Creativity Pre-test and IQ Covaried
(Grade Seven)[1]

Sub-test	Post-test Means		Adjusted Means		F	P
	Instructional	Control	Instructional	Control		
Asking Questions	7.05	7.32	7.01	7.36	<1	n.s.
Guessing Causes	7.24	5.11	7.21	5.15	6.420	<.05
Guessing Consequences	8.22	8.05	8.42	7.86	<1	n.s.
Unusual Uses	10.10	10.16	10.04	10.21	<1	n.s.
Product Improvement	9.34	8.39	9.10	8.61	<1	n.s.
Just Suppose	9.10	8.80	8.90	8.98	<1	n.s.
Fluency Sub-total	17.32	16.84	17.30	16.86	<1	n.s.
Flexibility Sub-total	17.15	17.48	17.03	17.59	<1	n.s.
Originality Sub-total	16.34	13.50	16.03	13.79	2.072	n.s.
Standardized Total	50.40	49.52	50.32	49.59	<1	n.s.

---

[1] Instructional group, N=41; Control group, N=44; df= 1,81

which there were no significant differences, the instructional group mean was greater than the control group mean in eight. (The absolute value of differences between the groups' post-test means and between the groups' adjusted post-test means tended to be quite small).

At the fifth-grade level (Tables 19b and 20b) there were no significant differences between the instructional group and control group adjusted means on any of the 20 comparisons made, although in 16 of the 20 comparisons the instructional group mean was greater than the control group mean.

At the sixth-grade level (Tables 19c and 20c), the control group mean was significantly greater than the instructional group mean for the Guessing Consequences subtest, when pre-test was statistically controlled (6.17 vs. 4.48, F=4.997, p<.05 with 1,70 df). This difference remained significant when both pre-test and IQ were statistically controlled (6.24 vs. 4.44, F=6.509, p<.05, with 1,69 df). There were no significant differences on any of the other comparisons. The instructional group mean was slightly greater than the control group mean on five of the remaining 18 comparisons.

At the seventh-grade level (Tables 10d and 20 d,) the instructional group mean was significantly greater than the control group mean on the Guessing Causes sub-test, when creativity pre-test was statistically controlled (7.20 vs.

5.15, F=5.787, $p<.05$ with 1,82 df). This difference remained significant when both IQ and creativity pre-test were statistically controlled (7.21 vs. 5.15, F=6.420, $p<.05$ with 1,81 df). Differences did not reach statistical significance on any of the other comparisons. In the remaining 18 comparisons, the instructional group adjusted mean was slightly greater than the control adjusted mean in ten comparisons.

General Problem Solving. The results of the General Problem Solving test are also presented by grade level. For each of the four grade levels, instructional and control group comparisons are presented for each of the five sets of scores obtained:

(a.) Analyses of proportion of pupils in each treatment solving correctly the Type C problems, using the Chi-square technique;

(b.) Analyses of the proportion of pupils in each treatment giving an acceptable response to the unsolvable problem;

(c.) Analyses of covariance, comparing mean number of solutions by instructional and control groups for each of the Type 0 problems, controlling statistically for the effects of IQ.

(d.) Analyses of covariance, comparing mean total number of Type C solutions for instructional and control groups, controlling statistically for the effects of IQ;

(e.) Analyses of covariance, comparing instructional and control group means on Standardized Type 0 Problem Total Scores, controlling statistically for the effects of IQ.

These results are summarized in Tables 21-25.

Table 21 summarizes the results for Type C problems by grade level and condition. At the fourth-grade level, very few pupils were able to solve the four problems. The Marble problem was solved by six of the 47 instructional group pupils, and by two of the 52 control pupils ($\chi^2=2.64$, $.10<p<.20$ with 1 df). The Water Jar problem was solved by ten control pupils, and by two instructional group pupils ($\chi^2=5.20$, $p<.05$ with 1 df). Ten instructed pupils and eleven control pupils solved the Candle problem ($\chi^2<1$,n.s.) and no one solved the Nine Dot problem.

At the fifth-grade level, 13 of the 44 instructional group pupils solved the Marble problem, compared with three of the 39 control pupils ($\chi^2=6.34$, $p<.02$ with 1 df). On each of the other three problems in this group, there were no significant differences between proportions of pupils solving the problems. Slightly more instructed pupils than control pupils solved the Candle Problem (18/44 compared with 13/39). Three pupils in each group solved the Water Jar problem and no one solved the Nine Dot problem.

At the sixth-grade level, there were no significant differences between treatment group proportions for any of the four problems. Nine instructed pupils and 12 control pupils solved the Marble problem. Twenty-one instructed pupils and 23 control pupils solved the Candle problem. Eleven instructed pupils and seven control pupils solved the Water Jar problem. No one solved the Nine Dot problem.

At the seventh-grade level, three control pupils and no instructed pupils solved the Nine Dot problem ($\chi^2=2.898$, $p<.10$ with 1 df). Differences between proportions for all other problems did not approach statistical significance. Fourteen instructional group pupils and 14 control pupils solved the Marble problem. Thirty-one instructed pupils and 28 control pupils solved the Candle problem. Nineteen instructed pupils and 18 control pupils solved the Water Jar problem.

With respect to total number of solutions (summed across all four type C problems), Table 22 indicates that there were no significant differences between instructional and control group means when IQ was statistically controlled. Examination of the means suggests that, especially in grades four, five, and six, pupils in both groups generally achieved very few solutions.

Table 23 summarizes the proportion of pupils in the instructional and control groups giving acceptable responses to the Unsolvable problem. At the sixth-grade level, a somewhat greater proportion of control pupils gave acceptable responses (13/44 vs. 6/49, $\chi^2=3.804$, $p<.10$ with 1 df). At each of the other three grade levels, there were no

Table 21:
General Problem Solving, Type C Problems
Proportion of Solutions by Grade and Condition

Problem	Instructional	Control	χ^2	P
Grade Four				
Marbles	6/47	2/52	2.64	.10<p<.20
Candles	10/47	11/52	<1	n.s.
Water Jar	2/47	10/52	5.20	<.05
Nine Dots	0	0	0	n.s.
Grade Five				
Marbles	13/44	3/39	6.34	<.02
Candles	18/44	13/39	<1	n.s.
Water Jar	3/44	3/39	<1	n.s.
Nine Dots	0	0	0	n.s.
Grade Six				
Marbles	9/49	12/46	<1	n.s.
Candles	21/49	23/46	<1	n.s.
Water Jar	11/49	7/46	<1	n.s.
Nine Dots	0	0	0	n.s.
Grade Seven				
Marbles	14/41	14/44	<1	n.s.
Candles	31/41	28/44	1.43	n.s.
Water Jar	19/41	18/44	<1	n.s.
Nine Dots	0	3/44	2.898	<.10

Table 23

General Problem Solving (Unsolvable Problem)
Proportion of Acceptable Responses By Grade and Condition

Grade	Proportion Giving Accepted Responses Instructional	Control	χ^2	P
Four	3/47	2/52	<1	n.s.
Five	10/44	5/39	1.370	n.s.
Six	6/49	13/46	3.804	<.10
Seven	21/41	19/44	<1	n.s.

At the fifth-grade level, the instructional group adjusted mean on Anagrams was significantly higher than the control group adjusted mean (2.46>1.76, F=7.686, p<.01). For the Antelope problem and the Rule Finding Problem, there were no significant differences between the groups' adjusted means. In both cases, the instructional group adjusted mean was slightly greater than the control group adjusted mean.

In both the sixth-and-seventh grades, there were no significant differences between the groups' adjusted means on any of the Type 0 problems. In the sixth-grade, control group adjusted means were slightly greater than instructional group adjusted means on each of the three problems. In the seventh-grade, the instructional group adjusted mean was greater than the control group adjusted mean only for the Antelopes problem.

Table 24:

General Problem Solving, Type O
Summary of Analyses of Covariance By Problems
Controlling For Effects of IQ,
By Grade and Condition

		Initial Means		Adjusted Means			
Grade	Problem	Instructional	Control	Instructional	Control	F	P
4	Antelopes[1]	6.19	9.13	6.01	9.30	9.829	<.01
	Anagrams[2]	2.13	1.33	2.09	1.36	10.085	<.01
	Rule Finding[3]	0.87	0.60	0.85	0.62	1.662	n.s.
5	Antelopes	6.77	7.08	7.02	6.80	<1	n.s.
	Anagrams	2.41	1.82	2.46	1.76	7.686	<.01
	Rule Finding	0.89	0.82	0.91	0.80	<1	n.s.
6	Antelopes	8.41	9.52	8.32	9.62	<1	n.s.
	Anagrams	2.08	2.09	2.05	2.12	<1	n.s.
	Rule Finding	1.10	1.15	1.07	1.18	<1	n.s.
7	Antelopes	21.00	18.84	21.12	18.73	<1	n.s.
	Anagrams	3.37	3.43	3.39	3.41	<1	n.s.
	Rule Finding	1.34	1.36	1.35	1.36	<1	n.s.

---

[1] Maximum Score Possible >60.

[2] Maximum Score Possible=10.

[3] Maximum Score Possible= 4.

Table 25 summarizes the analyses of covariance for Standardized Total Type 0 problem solving scores, comparing instructional and control groups by grade level. The total score was derived by standardizing (separately at each grade level) the distribution of scores on each of the three problems to $\bar{X}=50$ and S.D.=10, then summing the three standard scores and dividing by three. Table 25 reveals that in grades four, five, and seven, the instructional group mean total score was greater than the control group mean total score, when the effects of IQ were statistically controlled. In the sixth-grade, the control group adjusted mean was slightly greater than the instructional group adjusted mean. None of the differences reached statistical significance.

Arithmetic Problem Solving: Text Form. At each grade level, the mean total score of the instructional group was compared with the mean total score of the control group, statistically controlling for differences on arithmetic pre-test and IQ. Table 26 summarizes the results of these analyses.

None of the differences between instructional and control groups reached statistical significance. Control group adjusted means tended to be slightly higher than instructional group adjusted means at each of the four grade levels.

Arithmetic Problem Solving: Puzzle Form. At each grade level, two sets of analyses were conducted to compare

Table 25

General Problem Solving, Type O

Analysis of Covariance, Standardized Total Score, IQ Covaried, by Grade and Condition[1]

Grade[2]	Ttl. Score Means Instrctnl.	Ttl. Score Means Cntrl.	Adjusted Means Instrctnl.	Adjusted Means Cntrl.	F	P
4	50.72	49.32	50.43	49.59	<1	n.s.
5	50.73	49.23	51.05	48.86	3.025	n.s.
6	49.71	50.35	49.51	50.56	<1	n.s.
7	50.12	49.82	50.22	49.73	<1	n.s.

---

[1] To produce the Standardized Total Score, each problem was standardized to $\bar{x}=50$, S.D.=10, by grade level. Total score represents the arithmetic average of the standardized problem scores.

[2] Grade four: Instructional N=47, Control N=52 (1,96 df)
Grade five: Instructional N=44, Control N=39 (1,80 df)
Grade six: Instructional N=49, Control N=46 (1,92 df)
Grade seven: Instructional N=41, Control N=44 (1,82 df)

instructional and control group scores on the Puzzle Form of the Arithmetic Problem test. These were:

(a.) Chi-square analyses comparing the proportion of pupils in each treatment group correctly solving each problem;

(b.) Analysis of covariance, comparing total score means of treatment groups, controlling statistically for the effects of IQ and Arithmetic Pre-test.

These results are summarized in Tables 27 and 28.

Table 27 presents the results of comparisons of the proportion of pupils in both treatment groups (instructional

Table 26:
Arithmetic Problem Solving: Text Form,
Analysis of Covariance
Controlling for Pre-Test and IQ, by Grade and Condition

	Initial Means[1]		Adjusted Means			
Grade	Instrctnl.	Cntrl.	Instrctnl.	Cntrl.	F	P
Four	2.18	2.46	2.28	2.38	< 1	n.s.
Five	3.05	4.18	3.35	3.92	1.338	n.s.
Six	7.00	5.26	6.11	6.19	< 1	n.s.
Seven	7.52	8.62	7.96	8.36	1.014	n.s.

---

[1] Maximum Possible Score=16.

and control) who correctly solved each of the ten problems in the Puzzle form of the Arithmetic Problem Solving test. In grade four, there were no significant differences between groups. In grade five, the control group proportion exceeded the instructional group proportion only on problem 9 (3/18 vs. 0/25, $\chi^2=4.48$, $p<.05$ with 1 df). There were no other significant differences between groups at the fifth-grade level. At the sixth-grade level, there were no significant differences between groups. On problem 7, the instructional group mean was greater than the control with $\chi^2=2.68$, $.10<p<.20$, 1 df. At the seventh-grade level, there were no significant differences between treatment groups on any of the ten problems.

Table 28 presents the results of analyses of covariance for total scores on the Puzzle Form of the Arithmetic Problem Solving Test (an arithmetic sum of problems correctly solved)

Table 27

Arithmetic Problem Solving: Puzzle Form
Proportion of Pupils Solving Each Problem
By Grade Level and Condition

		Proportion Solving Correctly			
Grade - Problem		Instructional	Control	X^2	P
Four	1	0	0	0	n.s.
	2	0	0	0	n.s.
	3	8/24	11/26	<1	n.s.
	4	0	0	0	n.s.
	5	2/24	5/26	1.23	n.s.
	6	0	0	0	n.s.
	7	1/24	3.26	<1	n.s.
	8	3/24	1/26	1.27	n.s.
	9	1/24	1/26	<1	n.s.
	10	0	0	0	n.s.
Five	1	1/25	0/18	<1	n.s.
	2	0	0	0	n.s.
	3	10/25	4/18	1.51	n.s.
	4	2/25	1/18	<1	n.s.
	5	11/25	7/18	<1	n.s.
	6	2/25	3/18	<1	n.s.
	7	2/25	2/18	<1	n.s.
	8	3/25	3/18	<1	n.s.
	9	0/25	3/18	4.48	$p<.05$
	10	0	0	0	n.s.

-continued-

(Table 27:
continued)

Grade	Problem	Proportion Solving Correctly Instructional	Control	X^2	P
Six	1	0/24	1/22	1.12	n.s.
	2	0/24	1/22	1.12	n.s.
	3	2/24	2/22	<1	n.s.
	4	1/24	0/22	<1	n.s.
	5	12/24	12/22	<1	n.s.
	6	1/24	3/22	1.42	n.s.
	7	5/24	1/22	2.68	$.10<p<.20$
	8	11/24	6/22	1.42	n.s.
	9	2/24	1/22	<1	n.s.
	10	0	0	0	n.s.
Seven	1	0	0	0	n.s.
	2	0	0	0	n.s.
	3	5/17	8/23	<1	n.s.
	4	0	0	0	n.s.
	5	13/17	20/23	<1	n.s.
	6	4/17	3/23	<1	n.s.
	7	7/17	5/23	1.76	n.s.
	8	6/17	6/23	<1	n.s.
	9	2/17	5/23	<1	n.s.
	10	0/17	1/23	<1	n.s.

at each of the four grade levels, comparing mean total score of the instructional and the control groups. IQ and pre-test scores were statistically controlled.

Table 28 indicates that there were no differences between treatment groups at any of the four grade levels. The mean number of solutions achieved, in the highest scoring group, was only slightly greater than two (of a possible score of ten).

Table 28:
Arithmetic Problem Solving: Puzzle Form
Analysis of Covariance on Total Score (Controlling for IQ and Pre-test) by Grade and Condition

	Initial Means		Adjusted Means			
Grade	Instrctnl.	Cntrl.	Instrctnl.	Cntrl.	F	P
Four	0.63	0.81	0.64	0.79	<1	n.s.
Five	1.24	1.28	1.33	1.15	<1	n.s.
Six	1.42	1.41	1.22	1.62	1.305	n.s.
Seven	2.18	2.09	2.24	2.04	<1	n.s.

---

[1] Maximum possible score=10.

Make Up Problems. Table 29 presents the results of a one-way analysis of variance comparison of the scores of pupils in the instructional and control groups on the Make Up Problems test. The table indicates that there were no significant differences between treatment groups' adjusted mean scores at any of the four grade levels.

Table 29:
Make Up Problems Test Comparisons
By Grade and Condition

	Instructional			Control				
Grade	Mean	S.D.	N	Mean	S.D.	N	F(df)	P
Four	1.02	2.73	46	2.04	4.60	52	1.71(1, 96)	n.s.
Five	1.82	3.57	44	3.65	6.85	40	2.42(1, 82)	n.s.
Six	5.29	6.50	48	3.95	5.73	44	1.09(1, 90)	n.s.
Seven	8.74	12.34	39	13.20	12.72	45	2.64(1, 83)	n.s.

Attitude Inventory. The results from the *Childhood Attitude Inventory* are presented by grade level for each of three scores: first, one-way analysis of variance comparisons of treatment group means on Part I of the inventory; second, one-way analysis of variance comparisons of treatment group means on Part II of the inventory; and finally, one-way analysis of variance comparisons of treatment group means for Total Score on the Inventory (the sum of Part I and Part II for each pupil.) These data are summarized in Table 30.

At the fourth-grade level, the instructional group mean was significantly greater than the control group mean for Part I (17.02 vs. 14.75, F=10.64, p<.01, with 1,97 df). On Part II, there were no significant differences between the instructional and control groups. On Total Score, the instructional group mean was significantly greater than the control group mean (30.07 vs. 27.64, F=4.07, p<.05 with 1,88 df).

Table 30:

Pupil Attitude Inventory Comparisons
By Grade, Part of Inventory, and Condition

Grade-Section	Instrctl. Grp.			Control Group				
of Inventory[1]	Mean	S. D.	N.	Mean	S. D.	N	F(df)	P
4 - Part I	17.02	3.46	47	14.75	3.46	52	10.64(1,97)	<.01
Part II	12.84	4.01	45	12.64	3.30	45	<1	n.s.
Total	30.07	6.01	45	27.64	5.11	45	4.07(1,88)	<.05
5 - Part I	18.84	5.03	44	15.05	4.41	38	12.96(1,80)	<.01
Part II	14.19	3.76	36	12.59	3.09	32	3.63(1,66)	n.s.
Total	32.89	6.53	36	27.91	6.59	32	9.79(1,66)	<.01
6 - Part I	19.81	4.59	48	16.62	4.24	42	11.63(1,88)	<.01
Part II	13.44	3.14	48	13.84	3.30	37	<1	n.s.
Total	33.25	6.57	48	30.35	6.19	37	4.27(1,83)	<.05
7 - Part I	22.93	4.04	41	19.30	4.15	44	16.69(1,83)	<.01
Part II	12.08	4.10	40	12.05	3.89	44	<1	n.s.
Total	34.95	6.56	40	31.34	6.49	44	6.42(1,82)	<.025

---

[1] Maximum possible scores: Part I=30; Part II=22; Total=52.

The results for the sixth-grade also indicated that the instructional group mean was significantly greater than the control group mean only for part I and for Total Score. There was no significant difference between the groups' means on Part II. On Part I, the Instructional group mean was 19.81; the control mean was 16.62. The F=ratio of 11.63 was significant beyond the .01 level, with 1 and 88 df. For total score, the instructional group mean was 33.25 compared with the control group mean of 30.35 ($F=427$, $p<.05$, with 1,83 df).

The results for the seventh-grade groups maintained the same pattern as those at the other three grade levels. The instructional group mean was significantly greater than the control group mean on Part I (22.93 vs. 19.30, $F=16.69$, $p<.01$ with 1,83 df), and on Total Score (34.95 vs. 31.34, $F=6.42$, $p<.025$, with 1,82 df). There was no significant difference between treatment groups on Part II.

Internal Analyses. The results of the internal analyses will be reported in two general sections. First, for each grade level Pearson product moment correlation coefficients will be reported between each of the five factors of pupil responses (see Chapter Two) and each of the three creative problem solving criteria. These criteria are: General Problem Solving, Type C Standardized Total Score; General Problem Solving, Type C Total Score; and Post-Creativity Standardized Total Score.

Finally, for each grade level, and for each of the same three creative problem solving criteria, analyses of covariance were conducted for each response factor, to compare adjusted means of pupils high on the factor, low on the factor, and control pupils, with the effects of IQ statistically controlled. (For the post-test Creativity Total score criterion, the effects of pre-Creativity Total Score were also statistically controlled.)

Correlational Data. The indicated correlation coefficients were tested, using Fisher's Z-transformation, to determine whether, at each grade level and separately for each creative problem solving criterion, there were significant differences among the correlations with the five response factors. That is, at a specified grade level, and for a specified creative problem solving criterion, were there significant differences among the five response factor correlation coefficients?

Since there were twenty subjects at each grade level, the standard error of the difference between each pair of Z-coefficients could be determined for all pairs, substituting in the following formula (33, p. 241):

$$SE_{D_Z} = \sqrt{\frac{1}{N_1-3} + \frac{1}{N_2-3}}$$

where $N_1=20$ and $N_2=20$ for each pair of coefficients.

The result, the standard error of the difference between two Z-coefficients was

$$SE_{D_Z} = \sqrt{1/17 + 1/17}$$

$$= \sqrt{.05882 + .05882}$$

$$= \sqrt{.11764}$$

$$\approx .343$$

Since all groups consisted of N=20, it was then possible to determine the value of a difference necessary for significance at a given level for any pair of Z-coefficients in the present comparison. At the 5% level of confidence, with 18 df, the value of t required was 2.11.

Thus:

$$t = \frac{\text{Difference}}{SE_D}$$

$$2.11 = \frac{\text{Difference [needed]}}{.343}$$

$$2.11(.343) = \text{Difference [needed]}$$

$$.72 = \text{Difference [needed]}$$

Table 31 presents the Pearson product moment correlations and the appropriately equivalent Fisher Z-coefficient for each criterion, each factor, and each grade level. At each of the four grade levels there were no Z-coefficients which were as much as .72 greater than any of the other Z-coefficients for the same creative problem solving criterion. Thus, there were no pairs of Z-coefficients which were different from each other at an acceptable level of significance ($p<.05$).

Table 31:

Pearson Product Moment Correlations and Fisher Z-Coefficients Between Program Response Factors and Creativity and Problem Solving Criteria, by Grade Level

		Factor[2]									
		Mem.-Org.		Production		Reorganiz.		Judgment		Attitude	
Grade-Criterion[1]		r	Z	r	Z	r	Z	r	Z	r	Z
4	GPSO	.34	.35	.34	.35	.19	.19	.39	.41	.58	.66
	GPSC	.40	.42	.60	.69	.41	.44	.39	.41	.05	.05
	CREAT.	.61	.71	.66	.79	.45	.48	.75	.97	.51	.56
5	GPSO	.59	.68	.57	.65	.58	.66	.55	.62	.16	.16
	GPSC	.35	.37	.47	.51	.40	.42	.55	.62	-.04	-.04
	CREAT.	.50	.55	.43	.46	.42	.45	.45	.48	.07	.07
6	GPSO	.22	.22	.31	.32	.22	.22	.15	.15	-.09	-.09
	GPSC	.47	.51	.43	.46	.39	.41	.19	.19	.24	.24
	CREAT.	.43	.46	.37	.39	.20	.20	.46	.50	-.05	-.05
7	GPSO	.71	.89	.74	.95	.74	.95	.74	.95	.27	.28
	GPSC	.50	.55	.60	.69	.66	.79	.59	.68	.19	.19
	CREAT.	.67	.81	.76	1.00	.70	.87	.72	.91	.38	.40

---

[1] N=20 in each group.

[2] The pair of Z-coefficients, separately for each criterion, and each grade level, representing the greatest difference between factors is underlined. Difference needed for significance ($p<.05$)=.72.

Analyses of Covariance. The results of the analyses of covariance comparing pupils high on each factor, pupils low on each factor, and a sample of control pupils, for each of the three creative problem solving criteria and each of the five response factors, at each grade level, are summarized in Tables 32-46.

These data may be examined by response factors. For each response factor, results for all criteria at each grade level are summarized.

Memory-Organization. These data are summarized in Tables 32-34. There were no significant differences among the three creative problem solving criteria and at any of the four grade levels.

Table 32:
Analyses of Covariance
Type O Problem Solving
Memory-Organization Factor (IQ Covaried)

Grade	- Condition	Initial Mean	Adjusted Mean	F(df)	P
4	High	53.71	52.12	<1	n.s.
	Low	47.29	49.20		
	Control	50.60	50.38		
5	High	55.00	53.01	1.052	n.s.
	Low	45.57	48.97	(2,20)	
	Control	50.30	49.31		
6	High	51.00	48.87	1.927	n.s.
	Low	46.17	45.72	(2,18)	
	Control	48.90	50.45		
7	High	59.43	55.15	1.241	n.s.
	Low	47.57	52.77	(2,20)	
	Control	51.30	50.65		

Table 33:
Analyses of Covariance
Type C Problem Solving
Memory-Organization Factor (IQ Covaried)

Grade	Condition	Initial Mean	Adjusted Mean	F(df)	P
4	High	0.57	0.38	<1	n.s.
	Low	0.0	0.23		
	Control	0.30	0.27		
5	High	1.29	0.98	1.848	n.s.
	Low	0.71	1.24	(2,20)	
	Control	0.60	0.45		
6	High	1.67	1.48	1.061	n.s.
	Low	0.50	0.46	(2,18)	
	Control	0.90	1.03		
7	High	2.29	1.79	<1	n.s.
	Low	1.00	1.60		
	Control	2.10	2.03		

Table 34
Analyses of Covariance
Post-Creativity Total Score
Memory-Organization Factor
(IQ and Pre-Creativity Total Covaried)

Grade	Condition	Initial Mean	Adjusted Mean	F(df)	P
4	High	52.71	51.60	1.941	n.s.
	Low	44.86	46.30	(2,19)	
	Control	50.70	50.47		
5	High	51.57	50.58	<1	n.s.
	Low	47.00	48.85		
	Control	51.40	50.98		
6	High	51.17	50.57	1.893	n.s.
	Low	44.83	45.61	(2,17)	
	Control	50.20	50.09		
7	High	55.50	51.14	<1	n.s.
	Low	47.33	54.46		
	Control	52.30	50.64		

Production. These data are summarized in Tables 35-37. There were also no significant differences among the adjusted means for pupils high on production, pupils low on production, and control pupils. These results obtained for all creative problem solving criteria and at all grades.

Reorganization. These data are summarized in Tables 38-40. At the seventh-grade level, for Type C Problem Solving, there was a significant difference among the groups' adjusted means (2.84>1.98>1.16, F=3.635, p<.05 with 2,19 df). For all other comparisons for this factor, there were no significant differences among adjusted means.

Judgment. These data are summarized in Tables 41-43. There were no significant differences among adjusted means, at any grade level, for Type O Problem Solving or for Post-Creativity Total Score. For Type C Problem Solving, there were no significant differences among adjusted means, in grades four, six, and seven. In grade five, there was a statistically significant difference among adjusted means, favoring the pupils low on the Judgment factor (1.43>1.04>0.43, F=4.192, p<.05 with 2,21 df).

Attitude. These data are summarized in Tables 44-46. There were no significant differences among adjusted means at any grade level, nor for any of the creative problem solving criteria.

Of the sixty analyses of covariance, there were two cases in which F-ratios reached statistical significance.

Table 35

Analyses of Covariance
Type O Problem Solving
Production Factor
(IQ Covaried)

Grade - Condition		Initial Mean	Adjusted Mean	F(df)	P
4	High	53.71	52.45	<1	n.s.
	Low	47.71	49.19		
	Control	50.60	50.45		
5	High	54.29	52.04	<1	n.s.
	Low	47.14	50.51		
	Control	50.30	49.51		
6	High	51.50	48.83	<1	n.s.
	Low	46.83	47.21		
	Control	48.90	50.28		
7	High	57.71	54.40	<1	n.s.
	Low	45.33	50.40		
	Control	51.30	50.58		

Table 36

Analyses of Covariance
Type C Problem Solving
Production Factor
(IQ Covaried)

Grade - Condition		Initial Mean	Adjusted Mean	F(df)	P
4	High	0.86	0.72	1.204	n.s.
	Low	0.0	0.17	(2,20)	
	Control	0.30	0.28		
5	High	1.00	0.64	1.972	n.s.
	Low	0.71	1.26	(2,20)	
	Control	0.60	0.47		
6	High	1.67	1.39	1.027	n.s.
	Low	0.33	0.37	(2,18)	
	Control	0.90	1.04		
7	High	2.43	1.99	<1	n.s.
	Low	0.83	1.50		
	Control	2.10	2.01		

Table 37

Analyses of Covariance
Post-Creativity Total Score
Production Factor
(IQ and Pre-Creativity Total Covaried)

Grade - Control		Initial Mean	Adjusted Mean	F(df)	P
4	High	51.71	50.73	<1	n.s.
	Low	46.43	47.66		
	Control	50.70	50.53		
5	High	50.57	49.74	<1	n.s.
	Low	47.17	49.08		
	Control	51.40	50.83		
6	High	50.00	49.20	<1	n.s.
	Low	44.80	46.39		
	Control	50.20	49.88		
7	High	57.67	53.39	<1	n.s.
	Low	45.20	53.55		
	Control	52.30	50.69		

Table 38

Analyses of Covariance, Type O Problem Solving
Reorganization Factor
(IQ Covaried)

Grade - Condition		Initial Mean	Adjusted Mean	F(df)	P
4	High	53.57	52.29	< 1	n.s.
	Low	50.29	51.83		
	Control	50.60	50.41		
5	High	52.29	51.24	< 1	n.s.
	Low	46.29	51.25		
	Control	50.30	48.96		
6	High	51.00	47.92	< 1	n.s.
	Low	49.29	49.87		
	Control	48.90	50.34		
7	High	56.67	54.36	1.191	n.s.
	Low	46.29	50.74	(2,19)	
	Control	51.30	49.57		

Table 39
Analyses of Covariance
Type C Problem Solving
Reorganization Factor
(IQ Covaried)

Grade-Condition		Initial Mean	Adjusted Mean	F(df)	P
4	High	0.71	0.55	<1	n.s.
	Low	0.0	0.20		
	Control	0.30	0.28		
5	High	1.00	0.60	1.096	n.s.
	Low	0.43	1.08	(2,20)	
	Control	0.60	0.42		
6	High	1.50	1.15	<1	n.s.
	Low	0.57	0.64		
	Control	0.90	1.07		
7	High	3.00	2.84	3.635	<.05
	Low	0.86	1.16	(2,19)	
	Control	2.10	1.98		

Table 40
Analyses of Covariance
Post-Creativity Total Score
Reorganization Factor
(IQ and Pre-Creativity Total Covaried

Grade - Condition		Initial Mean	Adjusted Mean	F(df)	P
4	High	51.43	50.35	<1	n.s.
	Low	47.00	48.39		
	Control	50.70	50.48		
5	High	52.71	51.51	<1	n.s.
	Low	46.50	49.42		
	Control	51.49	50.49		
6	High	50.83	50.88	<1	n.s.
	Low	49.67	49.85		
	Control	50.20	50.06		
7	High	55.40	52.95	<1	n.s.
	Low	46.33	51.63		
	Control	52.30	50.35		

Table 41
Analyses of Covariance
Type O Problem Solving
Judgment Factor
(IQ Covaried)

Grade -	Condition	Initial Mean	Adjusted Mean	F(df)	P
4	High	53.71	51.92	<1	n.s.
	Low	47.86	50.10		
	Control	50.61	50.29		
5	High	55.57	51.59	<1	n.s.
	Low	48.25	53.01		
	Control	50.30	49.28		
6	High	51.14	49.68	<1	n.s.
	Low	47.33	46.87		
	Control	48.90	50.20		
7	High	58.83	55.21	1.221	n.s.
	Low	46.29	50.84	(2,19)	
	Control	51.30	50.29		

Table 42
Analyses of Covariance
Type C Problem Solving
Judgment Factor
(IQ Covaried)

Grade -	Condition	Initial Mean	Adjusted Mean	F(df)	P
4	High	0.57	0.33	<1	n.s.
	Low	0.14	0.44		
	Control	0.30	0.26		
5	High	1.71	1.04	4.192	<.05
	Low	0.63	1.43	(2,21)	
	Control	0.60	0.43		
6	High	1.43	1.28	<1	n.s.
	Low	0.67	0.62		
	Control	0.90	1.03		
7	High	2.33	1.80	<1	n.s.
	Low	0.86	1.53		
	Control	2.10	1.95		

Table 43

Analyses of Covariance

Post-Creativity Total Score

Judgment Factor

(IQ and Pre-Creativity Total Covaried)

Grade	- Condition	Initial Mean	Adjusted Mean	F(df)	P
4	High	52.71	51.65	1.254	n.s.
	Low	46.29	47.56	(2,19)	
	Control	50.70	50.55		
5	High	52.57	50.18	<1	n.s.
	Low	47.13	49.14		
	Control	51.40	51.46		
6	High	51.43	50.67	<1	n.s.
	Low	45.83	47.23		
	Control	50.20	49.89		
7	High	55.60	51.75	<1	n.s.
	Low	46.33	52.27		
	Control	52.30	50.66		

Table 44

Analyses of Covariance

Type 0 Problem Solving

Attitude Factor

(IQ Covaried)

Grade	- Condition	Initial Mean	Adjusted Mean	F(df)	P
4	High	54.14	53.06	1.294	n.s.
	Low	46.25	47.45	(2,21)	
	Control	50.60	50.40		
5	High	53.29	53.11	<1	n.s.
	Low	49.88	50.37		
	Control	50.30	50.03		
6	High	48.00	47.57	<1	n.s.
	Low	48.20	48.01		
	Control	48.90	49.29		
7	High	54.43	54.05	1.131	n.s.
	Low	48.00	50.98	(2,19)	
	Control	51.30	49.78		

Table 45
Analyses of Covariance
Type C Problem Solving
Attitude Factor
(IQ Covaried)

Grade -	Condition	Initial Mean	Adjusted Mean	F(df)	P
4	High	0.43	0.13	1.777	n.s.
	Low	0.38	0.70	(2,21)	
	Control	0.30	0.24		
5	High	1.00	0.97	1.838	n.s.
	Low	1.13	1.21	(2,21)	
	Control	0.60	0.55		
6	High	1.14	1.13	<1	n.s.
	Low	0.40	0.39		
	Control	0.90	0.91		
7	High	1.71	1.68	<1	n.s.
	Low	1.00	1.24		
	Control	2.10	1.98		

Table 46
Analyses of Covariance
Post-Creativity Total Score
Attitude Factor
(IQ and Pre-Creativity Total Covaried)

Grade -	Control	Initial Mean	Adjusted Mean	F(df)	P
4	High	51.43	49.78	<1	n.s.
	Low	46.13	48.13		
	Control	50.70	50.25		
5	High	51.43	52.25	1.420	n.s.
	Low	49.86	48.59	(2,19)	
	Control	51.40	51.71		
6	High	49.57	49.77	<1	n.s.
	Low	48.50	48.75		
	Control	50.20	49.96		
7	High	54.00	52.67	<1	n.s.
	Low	46.40	51.03		
	Control	52.30	50.78		

Post-hoc comparisons of means were conducted for each of these two analyses, using Duncan's New Multiple Range Test (29).

For the Judgment-Evaluation factor, Type C Problem Solving, at the fifth-grade level, the contrast revealed that only the two most extreme means differed significantly ($p<.05$). Thus, the mean for pupils low on the factor was significantly greater than for the control pupils. No other pairs differed significantly.

For the Reorganization factor, Type C Problem Solving, at the seventh-grade level, the contrast again revealed that only the most extreme mean difference reached significance ($p<.05$). Thus, the mean score of pupils high on the factor was significantly greater than the mean score of pupils low on the factor. Contrasts for all other pairs did not reach significance.

These results must be viewed with caution, however. The writer felt that, since only two of sixty F-ratios reached significance, the results may be merely artifacts of the number of analyses conducted.

Summary

Chapter Three has presented the results of the study. It was composed of three principal sections: comparisons of instructional and control pupils on pre-test variables; comparisons of instructional and control pupils on post-test variables; and analyses of program response factors, comparing creative problem solving criterion scores of pupils high on each response factor, low on each response factor, and a sub-sample of control pupils. In each section, results of analyses were reported separately for each of the four grade levels in the study.

CHAPTER IV

DISCUSSION AND CONCLUSIONS

This chapter presents discussion of the results of the study. It has been organized in two sections: the effects of the instructional treatment, by grade level, on verbal creativity scores, general problem solving criteria, and arithmetic problem solving criteria; and comparisons of a sub-sample of control pupils, pupils high on program response factors, and pupils low on program response factors, by grade level, with respect to three creative problem solving criteria.

Effects of instruction: Instructional-Control Comparisons

In this section, the results of comparisons of instructional and control pupils' scores on Verbal Creativity criteria, General Problem Solving criteria, Arithmetic Problem Solving criteria, the Make Up Problems Test, and the Childhood Attitude Inventory for Problem Solving will be discussed. The results for each of these criteria will be discussed separately, and the results presented for each grade level.

Verbal Creativity. Of 80 analyses conducted, there were only five significant differences between treatment groups. Of these, only three favored the instructional group pupils (both analyses for Guessing Causes at the seventh-grade level and Originality Sub-total at the fourth grade level). Since one might expect about four differences to be significant merely as an artifact of the number of analyses performed, it cannot be concluded with any confidence that the instructional materials influenced the pupils' verbal creativity scores to any appreciable extent. Although this finding does not offer support for the effectiveness of the materials, it is not inconsistent with previous results. Ripple and Dacey (56), for example, found similar results with eighth graders. Also. Olton et al (49) found a general lack of treatment effects with respect to similar criteria in a study of fifth-grade pupils.

General Problem Solving. Of 40 comparisons, there were again very few significant differences between instructional and control groups. The instructional group pupils performed significantly better than control pupils on only three of the comparisons: the Marbles Problem in the fifth-grade, and Anagrams in both grades four and five. Control pupils performed significantly better than instructional group pupils on two problems at the fourth-grade level (Water Jar Problem and Antelopes Problem) and on one problem at the seventh-grade level, although there the result was only of marginal significance (Dots problem, p.<10). For all other comparisons, involving both Type O and Type C problems, there were no significant differences between groups. Thus, there was extremely little evidence to support the effectiveness of the programed instructional materials, and little distinction between the Type O and Type C problems, despite the suggestion of Davis (25) that Type C problems should be more critically influenced by transfer than Type O problems.

Arithmetic Problem Solving-Text Form. Again, there were no significant differences between instructional and control group means, when IQ and arithmetic pre-test scores were statistically controlled, at any of the four grade levels. Although, upon examination of the content of this test, the problems seem to be content-valid with respect to arithmetic problem solving, it may be that the behavior required to solve such problems does not in fact correspond to the general skills and strategies which the program proposes to develop. However, given that these problems approximate those encountered by the elementary pupil in studying arithmetic, and given that the program proposes to teach general skills which will be of value in a wide range of content-specific situations, the criterion does not seem inappropriate. That is, while the test may be rigorous, it is not irrelevant. There is, from these data, no support for the assumption that the Productive Thinking Program effectively develops skills or abilities which will transfer to arithmetic problem solving tests of the kind used in this study.

Arithmetic Problem Solving - Puzzle Form. There were no significant differences between treatment groups with respect to any of the ten problems in this test at grades four, six, and seven. At the fifth-grade level, significantly more pupils in the control group solved problem number nine than in the instructional group. There were no significant differences between the treatment groups for any of the other nine problems. There were no differences between groups on mean total number of solutions, when IQ and Arithmetic Pre-test scores were statistically controlled; this result obtained at each of the four grade levels. Thus, there was no support from these data for the assertion of positive transfer from the instructional materials to problem solving of the kind presented in this test. Caution must be exercised in interpreting these results, however, because of the extreme difficulty of many of the problems (see Appendix D), and the low split-half reliability of the test.

Make Up Problems. There were no significant differences between treatment groups on Make Up Problems test scores, at any of the four grade levels. This test was included to provide at least one independent measure of an aspect of the problem solving process, often referred to as 'sensitivity to problems.' Previous research by Covington and Crutchfield (see Chapter One) suggested the superiority of instructed pupils over controls with respect to process criteria as well as number of actual solutions produced. In the present study, however, there was no evidence to support the effectiveness of the programed materials with respect to the Make Up Problems criterion measure.

Attitude Inventory. There were consistent significant differences between treatment groups with respect to the Pupil Attitude Inventory. On Part I, dealing with the pupil's attitude about problem solving, and on Total Score, the instructional group means were significantly greater than control group means at each of the four grade levels studied. There were no significant differences between instructional and control group means on Part II. These results lend support to the effectiveness of the Productive Thinking Program with respect to the development of pupils' attitudes towards creative thinking and problem solving.

Summary and Interpretation. The obvious generalization which emerges from these results is that there is very little evidence of the effectiveness of the Productive Thinking Program with respect to verbal creative thinking abilities or problem solving skills as here defined. This generalization appears to be valid with respect to each of the criteria used except the Pupil Attitude Inventory, and at each of the four grade levels studied. There was no evidence of the differential effectiveness of the programed materials that was proposed in Chapter One. Further, the generalization appears to be valid for the present data with respect to general problem solving measures (GPS, and both APS forms), verbal creativity scores, all of which are measures of products, as well as for the Make Up Problems test, which has been considered as one measure more directly process-related. Further, there was no support for the assumption presented by Covington and Crutchfield (16) that the programed materials develop skills and abilities which would transfer to subject-specific problem solving situations.

In seeking to understand these results it is, of course, initially necessary to examine the procedures which were followed. The question of the internal validity of the experimental design must be raised. Campbell and Stanley (11) have suggested eight considerations which are essential to internal validity in experimental and quasi-experimental design. These considerations are:

1. History. Events which occured between the first and second measurement of the dependent variable(s) may compete with the experimental treatment as causes of change in the dependent variable. In this study, subjects were randomly assigned, by school, to condition. There were no extreme differences among participating schools with respect to general curriculum, distribution of socio-economic levels, distribution of reading and intelligence scores, or sex distributions. These factors suggest that it was unlikely that students in any participating schools had experiences

which differentially affected their performance on the dependent variable measures. In addition, teachers were asked to report any departures from routine school experiences, which might have differentially influenced the results of the study; no such instances were reported.

2. Maturation. Changes in subjects' performance may be related to changes in age, development, hunger, fatigue, etc. In the present study, a relatively short period of time intervened between pre- and post-testing, and differences were the same for both instructional and control group pupils. It seems unlikely that these factors could have influenced the results of the study.

3. Testing. It is possible that taking a pre-test may have effects on post-test performance. Although in the short period of time which elapsed between pre- and post-testing, this variable may be relevant for the verbal creativity measures, it certainly does not bear importantly on the other dependent variable measurements. There is no logical or empirical reason to suspect that pre-tests on Intelligence, Verbal Creativity, or Arithmetic Skills differentially effect performance on the post-test measures used in this study. In addition, both instructional and control group pupils received the same test batteries, so differential effects of testing between treatment groups would be unlikely.

4. Instrumentation. Changes may occur during measurement in observers or scorers. Given the satisfactory reliability coefficients for scorers on the Verbal Creativity measures, this factor is unlikely to provide a satisfactory interpretation. Scoring of other dependent variable measures was done by the writers, carefully checked, and is considered to be reliable.

5. Selection. Even when subjects are randomly assigned to condition, there might be significant differences between the means of instructional and control groups on important variables. In Chapter Three, results of analyses of variance on Pre-Test variables indicated that the groups did differ with respect to some Pre-Test scores. These differences are unlikely to have greatly influenced the results of the study; statistical controls for the effects of Pre-Test variables were appropriately included in the analyses of the data.

6. Selection-maturation interaction. The combined effects of these variables may have counteracted actual gains from the instructional treatment. It has already been suggested that neither of these variables separately may be judged to have affected the results of the study; it is also unlikely that the interaction provides an acceptable interpretation.

7. Experimental Mortality. There may be a differential loss of subjects from pre-test to post-test, caused by some

variable important in the experiment. There was no evidence of differential loss of subjects from classes which might be attributed to any variable known to be related to the instructional treatment.

8. Statistical regression. When groups are selected on the basis of being in the extremes of a range of pretest scores, part of the reason subjects are found in these extremes is the fortuitous distribution of test error variance. Since pupils in the present study were not selected on such bases, and statistical controls were employed to equate groups with respect to initial measures, this factor cannot have affected the results of the study.

Inasmuch as none of these considerations of internal validity appears to provide an acceptable explanation of the results, other interpretations will be considered. It is important to note, however, that such alternatives must be considered quite speculative.

One possible interpretation, of course, is that the Productive Thinking Program may not be as effective in developing pupils' creative thinking and problem solving abilities as had been suggested by early research.

In considering this alternative, it is interesting to note two recent developments. These are:

(1) Recent research, conducted outside the Berkeley area and by independent researchers (although in consultation with the members of the staff of the Berkeley Creativity

Project) has yielded results considerably less impressive than had been reported in early research. In Chapter One, previous research with these materials was reviewed. Covington and Crutchfield (16) reported highly significant treatment effects. Ripple and Dacey (56) found generally negative results with respect to treatment effects on eighth graders' verbal creativity, and only limited effects on a problem solving criterion. Olton et al. (49) reported significant training effects for fifth-grade pupils, but of considerably smaller magnitude (and less extensive in scope) than in previous studies.

(2) In preparing the materials for commercial publication, the authors have considered it necessary to prepare a supplementary teacher's guide and a series of supplementary exercises.[1] While such materials do not imply the need for modification of the materials necessarily, Covington has discussed this possibility (15).

Other equally appealing alternative interpretations are readily available, however. First, it may be contended that only slight differences could reasonably have been found in any event. Several factors contribute to the plausibility of such an interpretation. The program was used as an entirely self-instructional sequence, with a minimum of teacher involvement. It has been pointed out in recent research (5, 49, 58) that teacher involvement in programed instruction may increase its effectiveness by as much as 50%. Further, the authors of the Productive Thinking Program have recently prepared a teacher's guide to supplement the instructional program. This guide was not available for use during the present study; in addition, since the materials were originally proposed as self-instructional, it was decided to use them entirely in that way to minimize differential effects among teachers. In addition, the program was used in the most concentrated possible sequence: one lesson per day for sixteen consecutive days. The program's authors have also recently proposed that the use of the materials should be extended over a longer period of time, ranging from four weeks (49) to as long as an eight week period.[2] These results may testify, then, more to the importance and impact of the teacher in classroom learning than to the ineffectiveness of the instructional materials.

---

[1] Dr. Robert Olton, personal communication, March 1968.

Consideration of several sociological variables which may be related to the development of pupils' creative thinking and problem solving abilities also points toward the potential impact of the teacher's role.

First, it has been proposed that there is a negative relationship between conformity and creativity (19). In the classroom, as in any social group, pressures toward conformity exist. Norms are developed which regulate the classroom behavior of all group members. It is likely that, in our schools, conforming behavior is expected and rewarded (positively sanctioned) to a greater extent than independent, creative behavior (34, 67). Thus, the pressure of the group towards conformity may restrict the development and expression of creative behavior among individuals. Of course, particularly at the upper elementary school levels, the teacher plays a role of extreme importance in determining the norms which will develop in her classroom, and in sanctioning pupil behavior (3).

---

[2] Dr. Robert Olton, personal communication, March 1968.

Further, Brookover and his associates (7, 8) have demonstrated that, even when the effects of IQ are statistically controlled, students' self-estimates of their ability are significantly related to actual educational achievement. In the tradition of G. H. Mead and C. H. Cooley, such formulations stress that one's self-concept is developed through interaction with "significant others", which in turn influences one's behavior. Presumably, at the elementary school level, the teacher is one of the critical significant others. That is, the pupil may formulate his self-concept of his own ability (in this case, specifically of his own creative thinking and problem solving ability) as a result of interaction and feedback from the teacher. The direction and extent of the teacher's acknowledgment (or sanctioning) of creative behavior may be an important influence on the development of his self-concept of ability and thus on his behavior.

Another limiting factor which must be considered in evaluating the negative results obtained is the difficulty of the criterion measures. Since no generally accepted criterion for problem solving abilities has been developed for elementary school pupils, it was necessary to construct measures for use in this study. Although considerable effort was devoted to the task of building appropriate instruments, the general level of successful solution of the problems used was poor. The difficulty of the criterion measures probably contributed to some extent to the absence of pronounced treatment effects. Since extremely difficult test items cannot lead to optimum discrimination indices (28), it may be suggested that the difficulty of the problem solving criteria used in this study may have led to a very conservative estimate of the impact of the materials on the pupils (i.e., masking such true differences as might have existed given less difficult criteria).

Another possibility which must be considered is that the distribution of teachers in the present sample is in some way biased so as to minimize the possible effectiveness of the programed materials. This factor could take the form of a generally high level of emphasis on creative thinking and problem solving in the classrooms which participated. Since Olton et al., (49) reported evidence to suggest that the programed materials were more effective in non-facilitative classroom environments, a generally high level among the classrooms in the present study would work against the effectiveness of the materials. There is no direct evidence to confirm or deny such a bias in the present sample, so that this factor cannot be confidently dismissed. The writers' estimate, based on observations among all participating teachers, is that there was no systematic bias among the participating teachers which might have differentially affected the program's potential effectiveness. The generally poor performance among pupils on the problem solving measures also suggests that there is opportunity for substantial improvement among pupils - that the lack of treatment effects cannot be accounted for by an overall superior level of performance.

It is also possible, as another speculative interpretation of these data, that it is unreasonable to expect any instructional program which constitutes only a small proportion of the pupil's educational experience, to have a significant impact on cognitive abilities as complex as though involved in creative thinking and problem solving. The complexity of these abilities, and the implications of teaching for such abilities, has been attested by Covington

(15). It does not appear likely to the writers that any instructional sequence which involves the pupil for so little time (one-sixth of a school day, for no more than two-twenty-fifths of the school year's actual instructional time) will lead to consequential cognitive growth.

Another possible interpretation of the data is also related to the criterion measures which were employed in the study. An important distinction was pointed out in Chapters I and II: the distinction between process and product measures of problem solving abilities. It was noted that the present study has deliberately employed product oriented criteria, in order to test to the fullest the potential effectiveness of the instructional materials. It does seem to be a plausible interpretation to suggest that the Productive Thinking Program may be more effective with respect to certain problem solving process criteria rather than the kind of product measures employed in this study. In a number of previous studies in which these materials were tested, problem solving process criteria were employed which were similar to the training materials in several respects. They were presented in booklet form (quite similar in make-up and appearance as well as format), guided the pupil through a problem solving episode, asked for responses in similar format to that of the training booklets. It is possible, therefore, that pupils' scores on these measures reflected the formation of a highly specific learning set (that is, that the pupils learned specifically how to solve problem episodes presented in a particular programed format), rather than complex general problem solving abilities. In a recent research report, Tuckman, Henkleman, O'Shaughnessy, and Cole (74) found that subjects were able to develop "search sets" (the strategy of seeking, and skill in finding, short cuts in problem solving). Subjects were trained in utilizing search sets in problem solving. However, as criterion problems became increasingly dissimilar from practice problems, trained subjects became "notably unsuccessful in finding the shortcut solution" (74, p. 68). The authors conclude that the strategy of search could be made to transfer more readily than the skill required to search successfully. The problems presented pupils in early research with the Productive Thinking Program, by virtue of their similarity to the training problems, may not have adequately probed the ability of trained pupils to successfully apply the skills developed in rigorous problem episodes. Although pupils in the present study were not directly tested with reference to the extent to which they actually learned the "thinking guideposts", there is some evidence in the present results which suggests pupils may have been influenced by the program even though they were unable to successfully utilize their training in a problem solving task. There were consistent significant differences favoring the instructed pupils on Part I of the Attitude Inventory, which deals with the pupil's attitude about creative problem solving, and on the Total Score of the Attitude Inventory. This finding suggests that pupils did learn to react favorably to creative problem solving. They may not have been able to be successful, however, at applying successfully the skills, abilities, and attitudes developed to the quite different problems presented in the criterion measures.

Comparisons by Pupil Program Response Factors

The "internal analyses" -- comparisons of pupils' creative problem scores by pupil response factors rather than the broader "instructed-not instructed" dichotomy also resulted in generally similar findings. Of 60 analyses of covariance (five factors, three criteria, four grade levels), there were only two significant differences. These differences (Type C Problem Solving, Judgment-Evaluation factor, in grade five; and, Type C Problem Solving, Judgment-Evaluation factor, in grade seven) have been considered by the writer as probable artifacts of the number of analyses conducted. This is particularly a plausible interpretation in that, for the same criterion variable and the same response factor, pupils in the High group exceeded pupils in the Low group in the seventh-grade, while in the fifth-grade, pupils in the Low group significantly exceeded only the Controls. In the absence of any logical explanation of such results, it would appear that they were chance occurences.

In addition, correlational data did not suggest that pupil performance on any of the five factors was differentially related to performance on any of the three creative problem solving criteria. These results were found at all four grade levels. Thus, there was no indication that, if the categorizations represented a valid cross-section of the instructional content of the program, there was any reason to believe the program developed any specific factors to a greater or lesser extent than others. Such inferences are, however, cautiously offered because of the exploratory nature of the investigation and the simplicity of the procedures employed.

Several problems warrant notice. First, the categorizations of pupil responses into "factors" were logically derived by two students of human problem solving. Although the factors with which we emerged seemed consistent with formulations of steps or stages in problem solving, there is no empirical support at all for the identity of the factors, nor for the association of particular pupil responses with proposed factors. For this reason, the results of the present study with regard to the "internal structure" of the programed materials should be regarded very cautiously.

It is also possible that the analysis of the written responses which pupils are required to make may not fully sample or representatively assess the abilities, skills, and attitudes developed in the instructional sequence. The child may develop many of these abilities without overt responses, if he is to develop them at all. Such influences, however, are not open to systematic examination when research is conducted in the classroom setting, and may not even be able to be examined in a more carefully controlled setting, given the present format of the materials. The possibility that the responses which a pupil makes may not representatively sample the instructional content of the program, or the pupil's learning experience, also is related to a practical concern. Although pupils have been carefully instructed not to look ahead in the program, it is very difficult to prevent a child from doing that if he so chooses. As a result, there can be no guarantee that the responses which were examined after collecting the answer booklets were indeed the original responses of the pupils in a learning experience.

Finally, it should be noted that the constraints of time and support required that these analyses be conducted using a relatively small sub-sample of the instructional and control groups. The possible effects of selection factors were attempted to be controlled, by equating the groups statistically with respect to IQ. It is possible, however, that differences among pupils on variables uncontrolled in the present research design might have been systematically related to the pupils' patterns of response to the program factors. The small size of the sample itself may have been a factor affecting the stability of the results. It is quite clear, especially in view of the findings of the broader instructional-control comparisons, that future research must be addressed to the problems of "what is taught" and "what is learned" in order adequately to assess the usefulness of these instructional materials. It also seems necessary to address subsequent research to these questions in order to deal effectively with the problems of supplementing the training booklets and directing teacher involvement.

Conclusions

In this section, each of the specific questions raised in Chapter One will be reviewed, with reference to the results of the study.

Questions One and Two dealt with the internal structure of the instructional program:

1. What abilities are developed in the program?

2. If some abilities are more effectively developed than others by the instructional materials, to what extent do the abilities taught bear differentially on pupils' problem solving performance?

Pupils' responses were used to examine what is taught in the program. Five factors were identified: Memory-Organization; Production; Reorganization; Judgment-Evaluation; and, Attitudes. With respect to verbal creativity and general problem solving criteria, however, none of these factors appeared to be differentially related to pupils' performance at any grade level in the study. In addition, when the effects of IQ were statistically controlled, there were generally no significant differences among pupils high or low on each factor and control pupils, at any grade level, and for all three creative problem solving criteria.

Questions three and four dealt with the practical value of the instructional materials:

3. At what grade level(s), if at all, may the instructional materials be recommended for classroom use?

4. To what extent, if at all, must the instructional materials be revised, modified, or supplemented by other activities for optimal effectiveness in the classroom?

The results of this study indicated that there was no evidence of the effectiveness of the materials, except with respect to pupil Attitude statements, at any of the four grade levels. This does not constitute, however, a general indictment of the potential value of the materials. It was indicated that the present study presented a much more demanding test of the materials than would be expected in classroom use. It is important to note, however, that the results of this study clearly indicate the need for further research and modification of procedures in using the materials. It must be concluded that, when used on an entirely self-instructional basis, with only minimal teacher involvement and tested with difficult, product-oriented problem solving criteria, there was very little effect on pupil performance. The potential effectiveness of the instructional materials, when used over a longer period of time, with active teacher involvement and supplementary work for pupils, warrants further examination.

Questions Five, Six, and Seven dealt with the specific effects of the program on verbal creativity, general problem solving criteria, and arithmetic problem solving:

5. To what extent, if at all, does the instructional treatment facilitate performance on tests of verbal creativity, when pre-test scores and IQ are statistically controlled?

There was no evidence in this study, that under the experimental conditions employed, the programed materials had any facilitating effect on pupils' verbal creativity scores, at any of the four grade levels studied.

6. To what extent, if at all, does what is learned or developed in the instructional treatment lead to positive transfer to general problem solving situations?

At all four grade levels the instructional treatment appears to have facilitated the expression of desirable attitudes about creative thinking and problem solving among instructed pupils. There was no evidence, however, of positive transfer from the programed instruction to a process criterion (the Make Up Problems Test) or to scores on the General Problem Solving Test (both Type O and Type C problems), at any grade level.

The writers have suggested, however, that the absence of pronounced training effects with respect to these criteria may have resulted from several factors, including the demanding use of the program, the difficulty of the problem solving tests, or the sharp contrast in emphasis between the product-oriented criteria and the process-oriented training materials. In addition, it was also suggested that it may be unreasonable to expect that any "instructional package" which involves such a small proportion of the pupil's educational experience will have any significant impact on the development of highly complex cognitive skills and abilities.

7. To what extent, if at all, does the learning from the instructional treatment lead to positive transfer to arithmetic problem solving?

There was no evidence of such positive transfer, at any grade level, with respect to the Arithmetic Problem Solving Test-Puzzle Form or the Arithmetic Problem Solving Test-Text Form. These results were undoubtedly influenced by the difficulty of the test items. Nevertheless, the tests did represent problems which were appropriate for pupils at the grade levels involved and thus constituted an appropriate, albeit difficult, test of the extent of the effects of the instruction. It may also be contended that the abilities required for solving the problems in these tests did not directly correspond with the abilities which the programed materials propose to develop. Such an argument is largely speculative, however, in the absence of any empirical indication of what abilities are actually developed by the program (and none was so revealed in the internal analyses in this study!) or any empirical indication of the abilities required to solve the problems in the tests. At least at one level, the test was appropriate: the tests presented problems which elementary school pupils could be expected to be given. Although Covington and Crutchfield (16) proposed that there would be positive transfer from the instructional program to a wide range of subject-matter problem solving tasks, no evidence supporting that assertion was found in this study. This conclusion, of course, is also limited by the factors previously discussed: the strict procedures with which the materials were used, the timing of presentation, and the difficulty of the problems. In addition, the reliability of the Puzzles Form was lower than desirable for experimental use.

Implications and Suggestions for Research

There are at least four implications of this study which the writers contend are important enough to warrant specification and to which future research should be addressed.

A. The interpretation of these data has suggested that the programed material's potential effectiveness may have been limited by several procedural factors. Further research seems necessary to explore the actual importance of these factors:

1. Rate of Presentation. The authors of the instructional materials have recommended that the program be used at a slower rate. Research must be addressed to investigate whether such modification leads to increased effectiveness.

2. Teacher Involvement. The modification of procedures to suggest teacher involvement rather than an entirely self-instructional presentation raises several questions to which research should be addressed. These questions include the examination of the effects of the instruction, with varying degrees of teacher involvement; examination of individual differences among teachers which relate to the effectiveness of the materials; and interactions of varying degrees of teacher involvement with rates of presentation, format of the materials, and extent of supplementary practice.

3. Supplementary Practice. Research may be addressed to examining the differential effects of the program under varying degrees and kinds of supplementary practice.

4. Format of Materials. With the increasing impact of technology on American Education, research may fruitfully be addressed to the effect of differing formats on the usefulness of the instructional materials. At least two immediately obvious modifications are (1) a multi-media presentation; and (2) programing the instructional materials for computer-based presentation.

While such questions are of immediate practical consequence, in determining the optimal mode and manner of presentation, they may also be of important consequence for psychological and educational researchers who are concerned with the broader question, "Can complex creative problem solving abilities be developed through direct instruction?".

B. Another important question has implications for the psychological study of transfer in relation to problem solving as well as immediate importance in evaluating the effectiveness of the Productive Thinking Program. The writers have suggested that the results of the present study may be accounted for, at least in part, by the emphasis on products rather than processes in the criterion measures in contrast to the greater emphasis on processes in the instructional materials. Tuckman, Henkelman, O'Shaughnessy, and Cole have expressed a similar concern:

> ...limited educational exposure to...problem solving approaches may induce students to adopt the strategy to search when confronted with transfer situations, but leave them lacking the skill to successfully apply this strategy (74, p. 68).

Thus, an important direction for further research would be to test the effectiveness of the instructional materials, using several problem solving criteria which vary in their degree of similarity to the instructional materials. Previous research by the program's authors has utilized problem solving criteria highly similar to the programed instructional materials. This study has used criteria quite dissimilar from the "training problems." The contrast of results underlines the need for systematic exploration of the differential effects of criteria varying along specified dimensions.

C. A third implication seeme to the writers to be very important, particularly in view of the results of this study: that is, extending systematic investigation of the internal structure of the program.

From a psychological point of view, this problem is of primary importance in the continuing investigation of instructing elementary school pupils in problem solving. There is a need to refine and modify the factors which have been identified in this study, and perhaps to identify new factors. Such research will be a challenging problem in view of the difficulties mentioned in this Chapter, but appears to be of extreme importance in assessing what is taught and what is learned. Further research in this area is also of importance from the more practical point of view of the educator. The assessment of the program's value and the specifications of modifications of its content seem to depend to a large extent on a more complete understanding of what it is that the program teaches or the pupil actually learns.

D. The fourth implication is derived from the writers' contention in Chapter Four that it may not be reasonable to expect any brief instructional program to have an extensive and permanent impact on the development of complex cognitive skills and abilities.

That creative thinking and problem solving abilities are extremely complex has been persuasively argued by a number of scholars (e.g., 13, 15, 41). The implications of the complexity of such abilities for testing (15), and for curriculum development (13), have been pointed out by one of the authors of the Productive Thinking Program. It is difficult, in these writers' opinion, to assume that pupils' cognitive development will be effectively facilitated by anything less than an extensive program for the reorganization of the whole instructional program of the elementary school. Such a program also demands training teachers who are better prepared to facilitate the development of pupils' problem solving skills, abilities, and attitudes. This will involve, necessarily, training teachers in several areas. Teachers must be prepared to utilize effectively materials such as the Productive Thinking Program, to develop their own materials and supplementary programs, to present subject-matter content more effectively, and to deal more effectively with pupils in classroom interaction, in order to realize the development of complex intellectual abilities. In short, there seems to be a need for "total learning environment" conducive to the development of problem solving abilities (72).

In addition to psychological and educational research, then, which is addressed to specific questions concerning the effectiveness of instructional programs such as the Productive Thinking Program, the writers contend that there is a need for ambitious, creative pilot programs, in which researchers, media-specialists, curriculum workers, and classroom teachers work together to develop a classroom environment and instructional program which will facilitate the development of pupils' problem solving abilities.

REFERENCES CITED

1. Ahmann, J. S., and Glock, M. D. Evaluating Pupil Growth. Boston: Allyn and Bacon, 1967.

2. Ausubel, David P. "Fostering creativity in the school." Paper presented at "How Children Learn," Centennial Symposium. Toronto, Ontario, Canada: Phi Delta Kappa and Ontario Institute for Studies in Education, February 25, 1967.

3. Backman, C. W., and Secord, P. A Social Psychological View of Education. New York: Harcourt, Brace, and World, 1968.

4. Bish, G. G. "A study of the relationships of intelligence, achievement, creativity, anxiety, and Confidence among intermediate grade pupils." Doctoral dissertation, George Washington University, 1964.

5. Blount, N. S., Klausmeier, H. J., Johnson, S. L., Frederick, W. C., and Ramsay, J. G. "The effectiveness of programed materials in English syntax and the relationship of selected variables to the learning of concepts." Technical Report from the Research and Development Center for Cognitive Learning. Madison: University of Wisconsin, 1967, No. 17.

6. Bowers, J. E. "A study of the relationships among measures of productive thinking, intelligence, and ninth grade achievement." Doctoral dissertation, University of Minnesota (Minneapolis), 1966.

7. Brookover, W., Paterson, A., and Thomas, S. Self-Concept of Ability and School Achievement. U. S. Office of Education, Cooperative Research Project No. 845 (East Lansing: Bureau of Research and Publications, Michigan State University, 1962).

8. Brookover, W., Thomas, S., and Paterson, A. "Self-concept of ability and school achievement." Sociology of Education, 1964, 37, 271-278.

9. Brunelle, Eugene A. Creativity and Problem Solving: 1965-66 Bibliography. Buffalo, New York: Creative Education Foundation, 1968.

10. Bruner, Jerome S. (Ed.) Learning About Learning: A conference Report. Washington, D. C: United States Government Printing Office, 1966. (Cooperative Research Monograph No. 15; OE-12019).

11. Campbell, D. T., and Stanley, J. C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNally, 1963.

12. Cicirelli, V. G. "Form of the relationship between creativity, IQ, and academic achievement." Journal of Educational Psychology, 1965, 56, 303-308.

13. Covington, M. V. "Productive thinking and a cognitive curriculum." Invited paper presented at Symposium: Studies of the Inquiry Process: Problems of Theory, Description, and Teaching, Washington, D. C., Annual Meeting of the American Psychological Association, September 1967.

14. Covington, M. V. "A childhood attitude inventory for problem solving." Berkeley: University of California, unpub. mimeo., 1967.

15. Covington, M. V. "New directions in the appraisal of creative thinking." Journal of Educational Measurement, 1968, in press.

16. Covington, M. V. and Crutchfield, R. S. "Facilitation of creative problem solving." Programed Instruction, 1965, 4, 3-5, 10.

17. Covington, M., Crutchfield, R., and Davies, L. The Productive Thinking Program; Series One: General Problem Solving. Berkeley, California: Educational Innovation, Inc., 1966.

18. Covington, M. V., Crutchfield, R. S., and Davies, L. Teacher's Guide for the Productive Thinking Program; Series One: General Problem Solving. Berkeley, California: Educational Innovation, 1967.

19. Crutchfield, R. S. "Conformity and creative thinking." In: Gruber, H., Terrell, G., and Wertheimer, M. (Eds.) Contemporary Approaches to Creative Thinking. New York: Atherton, 1962.

20. Crutchfield, R. S. "Instructing the individual in creative thinking." In Individualizing Instruction, Princeton, New Jersey: Educational Testing Service, 1965; Also in Mooney, R. and Razik, T. (Eds) Explorations in Creativity. New York: Harper and Row, 1967, pp. 196-206.

21. Crutchfield, R. S. "Recent findings on training for the facilitation of productive thinking." 1968, in preparation.

22. Crutchfield, R., and Covington, M. "Programed instruction and creativity." Programed Instruction, 1965, 4, 1-2, 8-10.

23. Cunnington, B. F., and Torrance, E. P. The Imagicraft series. Boston: Ginn and Company, 1966.

24. Dacey, John S. Programed Instruction in Creativity and Its Effects on Eighth Grade Pupils. Unpublished doctoral dissertation, Cornell University, Ithaca, New York. June 1967.

25. Davis, Gary A. "Current status of research and theory in human problem solving." Psychological Bulletin, 1966, 66, 36-54.

26. Duncan, C. P. "Recent research on human problem solving." Psychological Bulletin, 1959, 56, 397-429.

27. Duncker, K. "On problem solving." (Trans. by L. S. Lees) Psychological Monographs, 1945, 58, Whole No. 270.

28. Ebel, R. L. "Measurement and the teacher." Educational Leadership, 1962, 20, 20-24.

29. Edwards, A. L. Experimental Design in Psychological Research. New York: Holt, Rinehart, Winston, 1960.

30. Erickson, G. "The predictive validity of a battery of creative thinking tests and peer nominations among University of Minnesota high school seniors seven years later." Master's research paper, University of Minnesota, 1966; cited by Torrance, E. P. Torrance Tests of Creative Thinking: Norms and Technical Manual. Princeton, New Jersey: Personnel Press, 1966.

31. Fleming, E. S. and Weintraub, S. "Attitudinal rigidity as a measure of creativity in gifted children." Journal Of Educational Psychology, 1962, 53, 81-85.

32. Gagné, R. M. "Problem solving." In A. W. Melton (Ed.), Categories of Human Learning. New York: Academic Press, 1964. Pp. 297-317.

33. Garrett, H. E. *Statistics in Psychology and Education.* New York: David McKay, 1958.

34. Getzels, J. W. and Jackson, P. W. *Creativity and Intelligence*. New York: John Wiley, 1962.

35. Guenther, W. *The Analysis of Variance*. Englewood Cliffs, N. J.: Prentice Hall, 1964.

36. Guilford, J. P. "Creativity" *American Psychologist*, 1950, *5*, 444-454.

37. Guilford, J. P. "Three faces of intellect." *American Psychologist*, 1959, *14*, 469-479.

38. Guilford, J. P. *Fundamental Statistics for Psychology and Education*. New York: McGraw-Hill, 1965.

39. Guilford, J. P. "Intelligence, 1965 model." *American Psychologist*, 1966, *21*, 20-26.

40. Guilford, J. P. "Creativity: yesterday, today, and tomorrow." *Journal of Creative Behavior*, 1967, *1*, 3-14.

41. Guilford, J. P. *The Nature of Human Intelligence.* New York: McGraw-Hill, 1967.

42. Gulliksen, H. *Theory of Mental Tests.* New York: Wiley, 1950.

43. Long, B. H. and Henderson, E. H. "Opinion formation and creativity in elementary children." Newark, Del,: University of Newark, 1964, mimeo; cited by Torrance, E. P., *Torrance Tests of Creative Thinking: Norms and Technical Manual.* Princeton, N. J.: Personnel Press, 1966.

44. Lorge, I. and Thorndike, R. *The Lorge Thorndike Intelligence Tests: Technical Manual.* Boston: Houghton Mifflin, 1962.

45. Luchins, A. S. "Mechanization in problem solving: the effect of Einstellung." *Psychological Monographs*, 1942, *54*, Whole No. 248.

46. Maltzman, Irving. "On the training of originality." *Psychological Review*, 1960, *67*, 229-242.

47. Myers, R. and Torrance, E. P. Ideabooks. Boston: Ginn and Company, 1966.

48. Nelson, J. F. "The contruction of a scale of teacher judgment of pupil creativity." Master's research paper, University of Minnesota at Duluth, 1963; cited by Torrance, 1966.

49. Olton, Robert M., Wardrop, James L., Covington, M. V., Goodwin, W. L., Crutchfield, R. S., Klausmeier, H. J., and Ronda, T. The Development of Productive Thinking Skills in Fifth-Grade Children. Madison, Wisconsin: Center for Cognitive Learning, University of Wisconsin: Technical Report No. 34, 1967.

50. Osborn, Alex F. Applied Imagination. New York: Charles Scribner's Sons, 1957.

51. Parnes, Sidney J. "Do you really understand brainstorming?" In: Parnes, S. J., and Harding, H. H. (Eds.) A Source Book for Creative Thinking. New York: Charles Scribner's Sons, 1962. Pp. 283-290.

52. Parnes, Sidney J. "Creative potential and the educational experience." Buffalo, New York: Creative Education Foundation, Occasional Paper No. 2, 1967.

53. Raaheim, K. "Problem solving and awareness of the missing part." Scandinavian Journal of Psychology, 1962, 3, 129-131.

54. Random House Unabridged Dictionary of the English Language. New York: Random House, 1967.

55. Razik, T. Creativity Studies and Related Areas. Buffalo, New York: State University of New York (Creative Education Foundation), 1965.

56. Ripple, R. E. and Dacey, J. S. "The facilitation of problem solving and verbal creativity by exposure to programed instruction." Psychology in the Schools, 1967, 4, 240-245.

57. Ripple, R. E., Glock, M. D., and Millman, J. "Learner characteristics and instructional mode: a search for disordinal interactions." Journal of Educational Psychology, in press.

58. Ryan, F. L. "The effectiveness of teacher involvement in a programed social studies sequence." Paper presented at Annual Meeting, American Educational Research Association, Chicago, February 1968.

59. Sassenrath, J. M. "Manifest anxiety and transfer of complex learning." Psychological Reports, 1963, 12, 913-914.

60. Sassenrath, J. M. "Test anxiety, manifest anxiety, and concept learning without awareness." Psychological Reports, 1963, 12, 71-81.

61. Smith, James A. Setting Conditions for Creative Teaching in the Elementary School. Boston: Allyn and Bacon, 1966.

62. Spitzer, H. Practical Classroom Procedures for Enriching Arithmetic. Saint Louis: Webster, 1956.

63. Torrance, E. P. "Developing creative thinking through school experiences." In: Parnes, S. and Harding, H. (Eds.) A Source Book for Creative Thinking. New York: Charles Scribner's Sons, 1962. Pp. 31-47.

64. Torrance, E. P. Guiding Creative Talent. Englewood Cliffs, New Jersey: Prentice Hall, 1962.

65. Torrance, E. P. Education and the Creative Potential. Minneapolis, Minnesota: University of Minnesota Press, 1963.

66. Torrance, E. P. Constructive Behavior; Stress, Personality, and Mental Health. Belmont, California, Wadsworth, 1965

67. Torrance, E. P. Rewarding Creative Behavior. Englewood Cliffs, New Jersey: Prentice Hall, 1965

68. Torrance, E. P. Torrance Tests of Creative Thinking: Norms and Technical Manual. Princeton, New Jersey: Personnel Press, 1966.

69. Torrance, E. P. "Uniqueness and Creativeness: the school's role." Educational Leadership, 1967, 24, 493-496.

70. Torrance, E. P. and Gupta, R. Development and Evaluation of Recorded Programed Experiences in Creative Thinking in the Fourth Grade. Minneapolis: University of Minnesota, Bureau of Educational Research, 1964.

71. Torrance, E. P. and Myers, R. Teaching Gifted Elementary Pupils How to Do Research. Minneapolis: Perceptive Publishing, 1962. Cited by Torrance, 1966.

72. Treffinger, D. J. and Ferris, D. "Pilot proposal to train teachers for encouraging and developing pupils' creative thinking abilities." Lafayette, Indiana: Purdue University (mimeo), June, 1968.

73. Treffinger, D. J. and Ripple, R. E. "The interaction of sex of teacher and sex of pupil as a factor affecting teachers' rankings of pupil anxiety." Psychology in the Schools, 1967, 4, 331-335.

74. Tuckman, B., Henkelman, J., O'Shaughnessy, G., and Cole, M. B. "Induction and transfer of search sets." Journal of Educational Psychology, 1968, 59, 59-68.

75. Wallace, H. R. "Creative thinking: a factor in sales production." Vocational Guidance Quarterly, 1961, summer, 223-226; cited by Torrance, 1966.

76. Wallace, H. R. "Creative thinking: a factor in the performance of industrial salesmen." Doctoral dissertation, University of Minnesota, 1966; cited by Torrance, 1966.

77. Warner, W. L., Meeker, M. and Eels, D. Social Class in America. New York: Harper Torchbooks, 1960.

78. Weisberg, P. and Springer, K. J. "Environmental factors in creative function." Archives of General Psychiatry, 1961, 5, 554-564.

79. Yamamoto, K. "Creativity and sociometric choice among adolescents." Master's thesis, University of Minnesota, 1960; cited by Torrance, 1966.

80. Yamamoto, K. "A study of the relationships between creative thinking abilities of fifth grade teachers and academic achievement." Doctoral dissertation, University of Minnesota, 1962; cited by Torrance, 1966.

81. Yamamoto, K. "Creative writing and school achievement." School and Society, 1963, 91, 307-308.

82. Yamamoto, K. "Evaluation of some creativity measures in a high school with peer nominations as criteria." Journal of Psychology, 1964, 58, 285-293.

APPENDIX A

ARITHMETIC PRE-TEST

(WITH LEVELS OF ITEM DIFFICULTY AND DISCRIMINATION)

YOUR NAME ______________________________

GRADE ______________________________

SCHOOL ______________________________

Page One

DO ALL WORK ON THIS SHEET

1. Add:

$$\begin{array}{r}27\\42\\+\ 36\\\hline\end{array}$$

2. Subtract:

$$\begin{array}{r}438\\-216\\\hline\end{array}$$

3. Divide:

$$9\,\overline{)8\ 1\ 9}$$

4. Add:

$$\begin{array}{r}2342\\1321\\+\ 7114\\\hline\end{array}$$

5. Multiply:

$$\begin{array}{r}411\\\times\ \ 6\\\hline\end{array}$$

6. Subtract:

$$\begin{array}{r}736\\-478\\\hline\end{array}$$

7. Multiply:

$$\begin{array}{r}293\\\times\ 32\\\hline\end{array}$$

8. Divide:

$$6\,\overline{)516}$$

9. 1/4 of 36 equals what number?

10. George bought a ball that cost 43¢, a pencil that cost 10¢, and an eraser that cost 29¢. How much change did he receive from a dollar?

11. One yard equals ____________ inches?

12. Find the average of 15, 17, and 19.

13. Multiply:

$$\begin{array}{r}316\\\times\ 27\\\hline\end{array}$$

14. Add:

$$\begin{array}{r}317\\963\\874\\532\\+\ 441\\\hline\end{array}$$

15. Divide:

$$21\,\overline{)126}$$

16. How much will 16 apples cost at 6¢ each?

Page Two

DO ALL WORK ON THIS SHEET

17. Divide:

$6\overline{)9012}$

18. Subtract:

$$\begin{array}{r} 8726 \\ -5937 \\ \hline \end{array}$$

19. There are ________________ quarts in 3 gallons.

20. Divide:

$27\overline{)6102}$

21. Multiply:

$$\begin{array}{r} 534 \\ \times\ 407 \\ \hline \end{array}$$

22. Add:

$$\begin{array}{r} 7346 \\ 2971 \\ 41 \\ +\ \ 316 \\ \hline \end{array}$$

23. Add:

$$\begin{array}{r} \frac{1}{4} \\ +\ \frac{3}{4} \\ \hline \end{array}$$

24. Tell what numbers from 1 - 20 are prime numbers.

25. Subtract:

$$\begin{array}{r} 8000 \\ -6319 \\ \hline \end{array}$$

26. Divide:

$23\overline{)\ 3036}$

27. Add:

$$\begin{array}{r} 6\frac{3}{5} \\ 7\frac{1}{10} \\ \hline \end{array}$$

28. Multiply

$$\begin{array}{r} 3429 \\ \times\ \ 460 \\ \hline \end{array}$$

29. Subtract:

$$\begin{array}{r} 8\frac{2}{3} \\ -\ \ 5\frac{1}{2} \\ \hline \end{array}$$

30. What are the prime factors of 24?

31. A room measures 14 ft. by 26 ft. What is its area?

DO ALL WORK ON THIS SHEET

32. Divide:

$53\overline{)\,3267}$

33. Multiply:

$\frac{2}{3} \times \frac{5}{7} =$

34. Add:

$$\begin{array}{r} \frac{1}{6} \\ +7\,\frac{2}{3} \\ \hline \end{array}$$

35. Divide: $\frac{1}{6} \div \frac{3}{4} =$

36. The average of four numbers is 20. The first three numbers are 19, 15, and 32. Find the fourth number.

37. Arrange in order of their size beginning with the smallest: 5.04, .005, 5, 55.5, 5.005.

38. Multiply:

$3\,\frac{1}{3} \times 3\,\frac{1}{5}$

39. Add: 3.62, 10.91, 4.7, 19.641

40. Write the set of all even numbers from 1 - 21 inclusive.

41. Find the product of 82.7 and 3.2.

42. Elmer was travelling from Buffalo to New York City, a distance of 420 miles. After he had gone 240 miles, how far did he still have to go?

43. If □ equals any number, and △ equals any other number, tell whether this number sentence is true or false:

□ + △ = △ + □

Page Four

DO ALL WORK ON THIS SHEET

44. Divide: $.26\overline{).00962}$

45. Mr. Johnson earns \$8900 per year. He spends \$150 per month for rent. How much of his yearly income is left for other expenses?

46. Add:

$$\begin{array}{r} -\ 6 \\ +\ 16 \\ \hline \end{array}$$

47. Find the value of N if: $\frac{N}{7} = 6$

48. Subtract:

$$\begin{array}{rrr} 3\text{ hr.} & 40\text{ min.} & 25\text{ sec.} \\ -\ 2\text{ hr.} & 50\text{ min.} & 30\text{ sec.} \\ \hline \end{array}$$

49. Find 20% of 80.

50. Mary can solve 7 problems in 15 minutes. How long will it take her at this rate to solve 35 problems?

Item Difficulty and Discriminating Power (N=370)

Item	Difficulty[1]	Discrim.[2]	Item	Difficulty	Discrim.
1	91	0.184	26	29	0.825
2	95	0.053	27	27	0.825
3	65	0.851	28	14	0.368
4	90	0.149	29	21	0.675
5	82	0.526	30	8	0.289
6	76	0.518	31	15	0.430
7	49	0.825	32	23	0.702
8	52	0.930	33	25	0.579
9	50	0.851	34	25	0.798
10	51	0.772	35	6	0.175
11	86	0.342	36	6	0.184
12	39	0.868	37	7	0.254
13	42	0.825	38	5	0.175
14	64	0.474	39	9	0.281
15	43	0.904	40	6	0.184
16	63	0.763	41	3	0.114
17	43	0.886	42	36	0.439
18	57	0.614	43	30	0.447
19	66	0.667	44	1	0.035
20	25	0.763	45	9	0.212
21	30	0.684	46	6	0.079
22	71	0.509	47	6	0.175
23	55	0.895	48	4	0.123
24	10	0.316	49	2	0.079
25	58	0.719	50	10	0.167

---

[1] Percent Responding Correctly.

[2] Upper and Lower 27%; Following Ahmann and Glock (1967, pp. 187-189), range +1.00-(-1.00).

APPENDIX B

GENERAL PROBLEM SOLVING TEST

Name ______________________ Teacher ______________________

Grade ______________________ Date ______________________

SOLVING PROBLEMS

1. DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO.
2. In this booklet you will find several problems to solve. Some of them may be very easy; others may be very difficult. You are not expected to be able to solve all of them correctly.
3. Write your answer in the space provided. Mark it "Ans." so we will know it is your answer!
4. If you find more than one good solution for a problem, you may give more than one answer. Label each answer.
5. Do all your work right in the test booklet. Do not use any other paper!
6. Try to answer every problem. DO NOT WASTE TOO MUCH TIME WITH A PROBLEM YOU ARE UNABLE TO SOLVE.
7. Now read silently the special directions for this booklet:

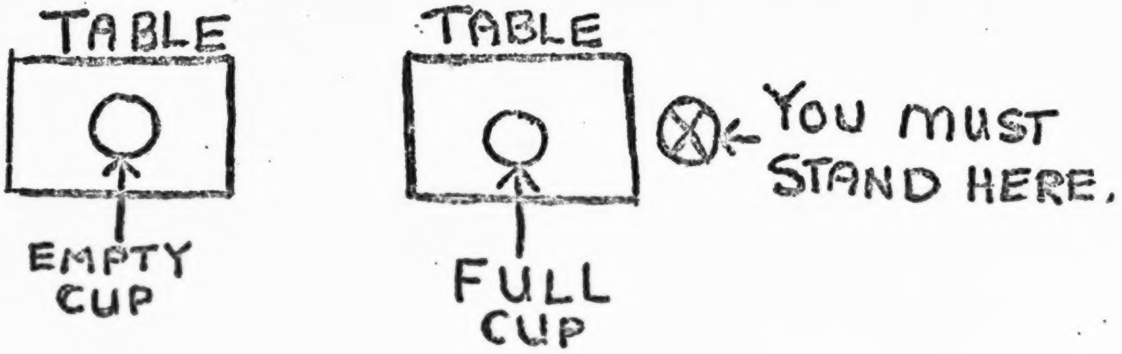
   A. Read each problem carefully. Do all your work in this booklet. Do not spend too much time on any problem that you find very difficult. Remember to label your answers.

   B. Some problems in this book may not have any solutions. If you think it is impossible to solve a problem, write "Can't be done." in the answer space, and tell why you think it can't be done. You may use brief sentences or drawings to help you.

1. In a room, there are two tables, six feet apart. On each table, there is a cup. In one cup, there are several marbles.

Problem:

Can you find a way to get the marbles from one cup into the other? You must stand behind the table on which the cup containing the marbles is now. (This is so you can't just carry the marbles to the other cup.) You may not move the tables or the cups. You may not throw the marbles. In the room are a scissors, a piece of string 3 feet long, and three newspapers. You may use these if you wish.



[Turn to page 2]

2.

Make as many correct English words as you can from the letters in this word:

ANTELOPES

(You may use each letter only as often as it appears in the word above, and you may not use any other letters. Proper names do not count!)

[Turn to page 3]

3.

Here are some words whose letters have been all mixed up!

For each one, rearrange the letters, if you can, to make one or more common words. You must use all the letters for each word.

(A) ETOSV

(B) DAERB

(C) EQITU

(D) ULEVA

(E) IWHGE

[Turn to page 4]

4.

John has 7 pockets and twenty dimes. He wants to put his dimes into his pockets so that each pocket contains a different number of dimes. How can he do this?

[Turn to page 5]

5. On a table, you find the following objects:

a fork;
a sheet of thin paper;
several thumbtacks;
three boxes of wooden matches;
a book;
three candles

On the wall, there is a large bulletin board.

Problem:

Find as many ways as you can to mount the 3 candles on the bulletin board, high enough so that you can look straight at them. You may only use the materials that are on the table to help you mount the candles!

[Turn to page 6]

6. A. Let's see if you can figure out how to play this game. Every time you say a word, I tell you a number that goes with the word.

Suppose you say "hospital." Then I say, "7." Here are some other words, with "their numbers":

It = 1

Dragon = 5

Good = 3

See if you can write the correct numbers for these words:

Development = _____

Morning = _____

B. Now, I'll change the rule I used to get the numbers. Now:

Hospital = 4

Good = 8

Dragon = 6

It = 10

What would be the new numbers for:

Development = _____

Morning = _____

[Turn to page 7]

7. At camp, it is your job to get the water. You need to get the correct amount, but the only way you have to measure how much water you have is by knowing how much your water cans will hold.

Problems: (1) You have 3 water cans. One holds 127 gallons. The second one holds 21 gallons. The third one holds 3 gallons. Knowing only these facts, how can you bring exactly 100 gallons of water in one trip from the water pump?

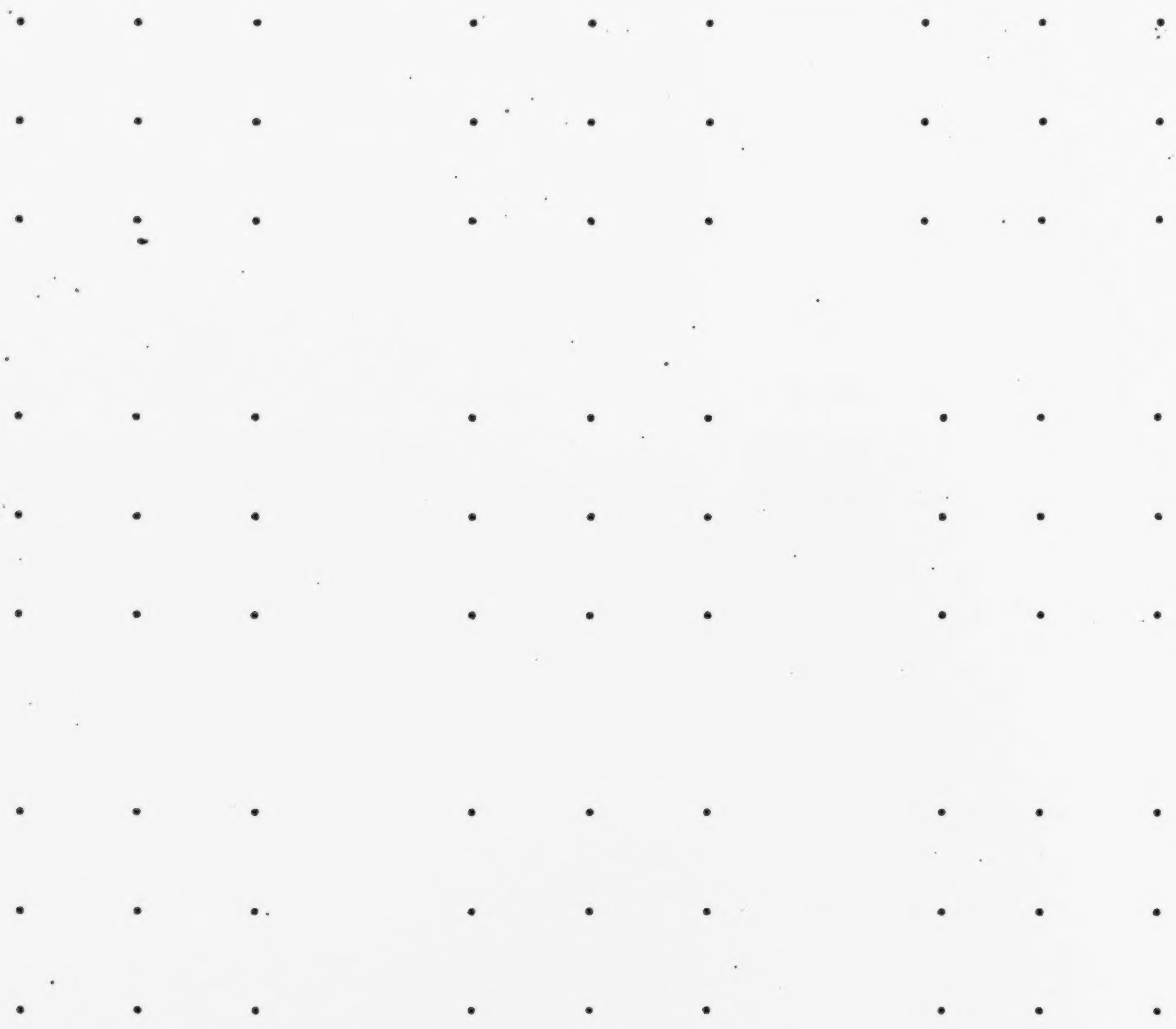
(2) You have 3 cans. One holds 140 gallons. One holds 20 gallons. The third holds 10 gallons. Knowing only these facts, how can you measure out exactly 90 gallons to bring back?

[Turn to page 8]

8.

Can you connect the 9 dots below with 4 straight lines, without lifting your pencil from the paper or drawing over any line a second time?

(Several groups of 9 dots have been drawn in case you need to start over. Make sure you label your final answer.)



[Stop Here.]

APPENDIX C

ARITHMETIC PROBLEM SOLVING TEST - TEXT FORM

(WITH LEVELS OF ITEM DISCRIMINATION AND DIFFICULTY)

Name_________________________ Teacher_______________________

Grade________________________ Date _______________________

SOLVING PROBLEMS

1. DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO.
2. In this b-oklet you will find several problems to solve. Some of them may be very easy, others may be very difficult. You are not expected to be able to solve all of them correctly.
3. Write your answer in the space provided. Mark it "ans." so we will know it is your answer!
4. If you find more than one good solution for a problem, you may give more than one answer. Label each answer.
5. Do all your work right in the test booklet. Do not use any other paper!
6. Try to answer every problem. DO NOT WASTE TOO MUCH TIME WITH A PROBLEM YOU ARE UNABLE TO SOLVE.
7. Now read silently the special directions for this booklet:

   A. These problems are about arithmetic. Do as many of them as you can. Remember to read each one very carefully, and to work only in this booklet. Label your answers. Do not spend too much time on a problem you don't know how to do.

(1) Connie's mother gave her eight pennies. Her father gave her one nickel and one dime. How much money did she have?

(2) A store made sales to 248 people in the morning and to 346 people in the afternoon. To how many people were sales made in the store that day? How many more sales were there in the afternoon than in the morning?

(3) Using only whole numbers, make this sentence true in as many ways as possible:

$$\triangle \times \square = 24$$

(4) There are 5 buses to take 250 pupils to a football game. If each bus will carry the same number of pupils, how many will there be in each bus?

(5) In one day, eight jet planes left an airport. Each jet plane carried 128 passengers. How many passengers left the airport by jet plane that day?

(6) A man borrowed $520. He agreed to pay it back in 8 equal payments. How much was each payment?

(7) A telephone book contains 92 pages of names. If, on the average, a page contains 340 names, about how many names are there in the telephone book?

[Turn to page 3]

(8) Paul figured that the wheels on his father's car turn around 829 times in traveling one mile. How many times will the wheels turn around on a 256-mile trip?

(9) A man gave $30 to three boys. Each boy received part of the money. Tom got one-half of the money. Fred got one-third of the money. Jim got one-sixth of the money. How much did each boy receive?

[Turn to page 4]

(10) A tower is 1248 feet high, with three braces every 12 feet.

a. How many braces are there?

b. If each brace weighs 9 ounces, what is the total weight of the braces?

(11) The ratio of John's weight to Bill's weight is 5:4. If Bill weighs 104 pounds, how much does John weigh?

[Turn to page 5]

(12) 65 is 65% of what number?

(13) If I add 8 to 3 times my age, the result is 50. How old am I?

(14) A boy is seven years older than his sister. In two years, he will be twice as old as his sister. How old is each person now?

[Turn to page 6]

(15) A collection of dimes and nickels is worth 80 cents. There are twice as many nickels as dimes. Find the number of dimes.

(16) Pretend that cows cost $5 each, pigs cost $10 each, and chickens cost $1 each. A man must spend exactly $50. He must buy at least ten animals, and he must buy some of each kind of animal. He may not buy ten or more of any one kind of animal, nor is it possible to have a fraction of an animal. How many of each kind of animal can he buy? [Can you find more than one solution?]

[Stop here.]

Item Difficulty and Discriminating Power (N=179)

Item	Difficulty[1]	Discrimination[2]
1	82	0.396
2	44	0.771
3	32	0.688
4	68	0.833
5	53	0.729
6	47	0.896
7	37	0.771
8	22	0.646
9	28	0.667
10	02	0.104
11	01	0.063
12	21	0.480
13	15	0.417
14	02	0.104
15	27	0.542
16	28	0.646

---

[1] Percent responding correctly.

[2] Upper 27% vs. Lower 27% (Ahmann and Glock, 1967).

APPENDIX D

ARITHMETIC PROBLEM SOLVING TEST: PUZZLE FORM

(WITH LEVELS OF ITEM DISCRIMINATION AND DIFFICULTY)

Name________________________ Teacher______________________

Grade________________________ Date _______________________

SOLVING PROBLEMS

1. DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO.

2. In this booklet you will find several problems to solve. Some of them may be very easy; others may be very difficult. You are not expected to be able to solve all of them correctly.

3. Write your answer in the spaces provided. Mark it "Ans." so we will know it is your answer!

4. If you find more than one good solution for a problem, you may give more than one answer. Label each answer.

5. Do all your work right in the test booklet. Do not use any other paper!

6. Try to answer every problem. DO NOT WASTE TOO MUCH TIME WITH A PROBLEM YOU ARE UNABLE TO SOLVE.

7. Now read silently the special directions for this booklet:

   A. These problems are about arithmetic. Do as many of them as you can. Remember to read each one very carefully, and to work only in this booklet. Label your answers. Do not spend too much time on a problem you don't know how to do.

1.

If a bottle and a cork together cost $1.10, but the bottle costs one dollar more than the cork, how much does the cork cost?

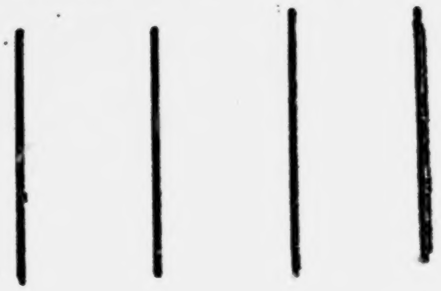
2.

A man said he could not make change for a dollar for me. He also said he could not give me change for a half-dollar. He couldn't even give me change for a quarter! I look at his change, and it added up to one dollar and fifteen cents. But all of the things he told me were true! How could this be? (There were no bills; the money was all American coins, and he did not have a silver dollar!)

[Turn to page 2]

3.

Look at the four lines below. Can you add exactly six more lines, and by doing that, make a name for the number five?



4.

Can you discover when half of 13 is 8? If so, show how:

5.

We have taken some of the numbers out of this addition problem, and written letters instead. Can you figure out what numbers should be put back in the example, instead of the letters, so it will add up correctly? (Each letter stands for a different number. 0 is the numeral zero, not the letter o.)

$$\begin{array}{r} 800W \\ 79X6 \\ +\ Y430 \\ \hline DOZ41 \end{array}$$

[Turn to page 3]

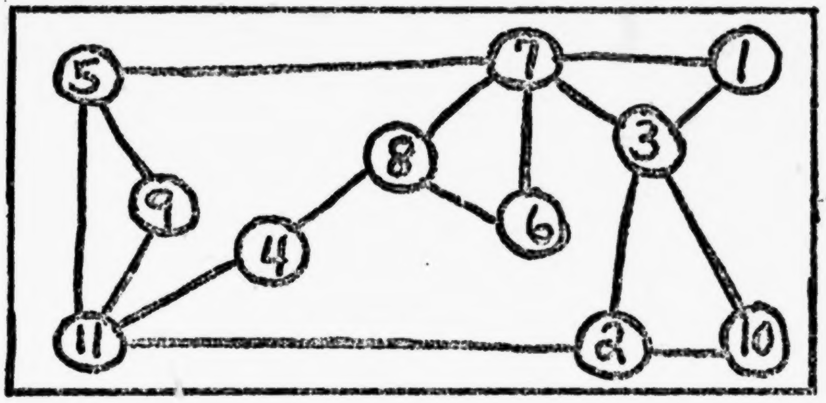
6.

Two sons and two fathers went hunting. They killed three rabbits. Then, they divided the rabbits equally - without cutting a rabbit. (No fractions, no remainders) How was this possible?

7.

Look at the buttons on the magic box. If you "push them" in the right order, you will solve this problem.

You must begin with Button 10 and end with Button 11. You can only go from one button to another if the two are connected by lines. The sum of the numbers on all the buttons you push must be 50.



I would push:

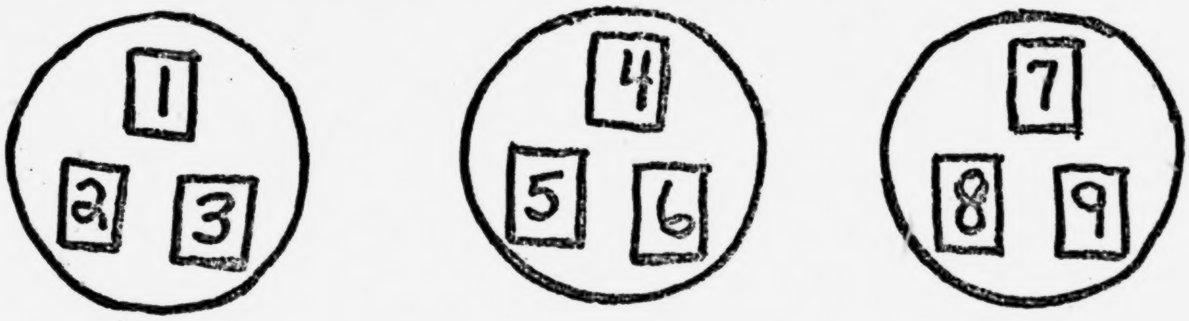
10, ___, ___, ___, ___, ___, ___, ___, ___, ___, 11 .

You don't have to use all the buttons.

[Turn to page 4]

8.

Can you move one card from one group to another so that the sum of the numbers in each group will then be equal. (Use an arrow with a line, and a brief sentence.)



9.

Suppose you know that in your drawer there are only three colors of socks - white, yellow, and blue. Pretend you have to find a pair of matching socks in the dark. What is the least number of socks you could pull from the drawer and be certain to have at least one pair of socks the same color?

10.

A group of hunters returned to their cabin with what they had shot. One man gathered all the squirrels and pheasants together. Another man asked him, "How many do we have?" The man answered, "We have 78 feet and 27 heads." How many squirrels and how many pheasants did the hunters get?

[Stop]

Item Difficulty and Discriminating Power (N=180)

Item	Difficulty[1]	Discrimination[2]
1	01.0	0.019
2	00.5	0.019
3	27.9	0.593
4	2.0	0.056
5	44.8	0.870
6	8.5	0.222
7	12.9	0.407
8	23.9	0.537
9	8.0	0.204
10	2.0	0.037

---

[1] Percent responding correctly

[2] Upper 27% vs. Lower 27% (Ahmann and Glock, 1967).

APPENDIX E

MAKE UP PROBLEMS TEST*

---

*Getzels, J. and Jackson, P. Creativity and Intelligence. New York: John Wiley, 1962. Used by permission of the authors and publisher.

Name____________________Grade____________________ Date______

MAKE-UP PROBLEMS

DIRECTIONS: Inside the booklet you will find a series of arithmetic problems. These problems are different from the ones that you are accustomed to doing because they contain more information than you need to get a single answer. In fact, the object of the test is to see how many different problems you can think of which might be solved with the information you are given. For example, you might be given the following material:

> John makes extra money in the spring and summer mowing lawns for people in his neighborhood. Sometimes he has more lawns to mow than he can handle. At such times he usually gets help from his two younger brothers, Fred and Tom. Fred can average half as much work as John in the same time, and Tom can average a third as much as John. John is paid by the job, and the people who hire him to not mind how many others help so long as John pays them out of his earnings and sees that the job is done right. Mrs. Jones pays John $9.00 a week to mow her lawn from the end of May until the middle of September, and John finds that he and his brothers can do the job in an hour and a half. Other neighbors pay John at the same rate as Mrs. Jones. Given this information, set up as many problems as you can involving John's working experience in the spring and summer, and that of his brothers.

In the space provided at the end of the paragraph, you might write any of the following problems, all of which might actually be worked knowing the information given. DO NOT WRITE DOWN ANY PROBLEMS WHICH CHOULD NOT BE WORKED WITHOUT ADDITIONAL INFORMATION.

1. What is a fair rate of pay for Fred and Tom?
2. How many lawns could the boys care for in an eight hour day?
3. If the boys work 10 hours every Saturday, how much money would they earn from May to September?
4. How much faster does Fred work than Tom?

And so on. Note that you would not ask a question such as, "How much money should John get for contacting neighbors and arranging to cut their lawns?" since this could not be answered from the information given.

You should give as many problems as you can for each paragraph of information. Try to give some problems for every paragraph rather than spending most of your time on just one or two paragraphs.

1. Mr. Smith decided to purchase a house whose cost was $15,000. He made a down payment of $5000, and agreed to pay the rest with monthly payments. Each monthly payment included a portion of the principal, an interest charge computed at the rate of 5 per cent per year, plus a charge for insurance which cost $129.50 per year. Mr. Smith found by talking to the former owner that it cost an average of $20 per month to heat the house. After he had owned the house for two years, he received $3000 through an uncle's will which he applied to what he still owed on the house. A year later he purchased a new stove and refrigerator on time payments which added $35 a month to his expenses. At the same time he added insulation to the house which cost him an additional $30 a month for 18 months, but which the contractor who installed it guaranteed would reduce his heating costs by 15 per cent. Given this information, set up as many problems as you can involving Mr. Smith's expenses in connection with the purchase and operation of his home.

2. The Park District of New City installs a swimming pool with a total capacity of 20,000 cubic feet. To fill the pool, two inlets with a potential of 20 and 10 cubic feet per minute respectively are available. A single drain at the deep end of the pool will remove water at the rate of 25 cubic feet per minute. A circulating pump is provided which moves the water in the pool through a filtration system at the rate of 5 cubic feet per minute. When the filters become clogged and require cleaning, the pool attendant is instructed to open the drain half way and then to open the larger inlet valve just enough so that the level of water in the pool remains constant. When the pool is to be cleaned, as it is once every week, the water is drained and the sides of the pool scrubbed. The draining and scrubbing together require 15 hours. Given this information, set up as many problems as you can concerning the operation of the pool.

3. Jack and Phil are writing a paper for their science class about falling bodies. They conduct some experiments by throwing rocks from a cliff to see how far they will fall in a certain time. When the rocks are dropped straight down they obtain the following rate:

Distance the Rock Falls	Time Required to Fall This Distance
4 feet	1/2 second
16 feet	1 second
64 feet	2 seconds
144 feet	3 seconds

They notice that when they throw the rocks straight out on a level the rocks fall just as fast and hit the ground in the same total time, 4.5 seconds, as when they were simply dropped. When they throw the rocks into the air just high enough to go over the top of a certain three standing at the edge of the cliff, it takes 9.5 seconds for the rock to fall to the bottom of the cliff. Given this information, set up as many problems as you can about Jack and Phil's experiment.

4. Mark and George are making a survey for a problem in their social science class. They want to find out how much money people who live in the neighborhood right around the school pay for rent. They decide to ask everyone who lives in the eight blocks surrounding the school how much they pay. Mark goes to all of the houses in four of the blocks and George to all the houses in the other four blocks. Of the 40 houses that Mark visits he gets answers from only 22 tenants; 14 tenants are away from home and 4 refuse to answer his questions. George visits 34 houses, getting answers from 18 tenants. Three tenants refuse to answer George's questions and 13 are not at home. Mark finds that the average rent paid by his 22 tenants is \$85 per month, with the highest being \$135 and the lowest \$47.50. George finds the average is \$97.50 with the highest being \$155 and lowest \$75.00.

   Since all of Mark's tenants live in apartment buildings managed by real estate companies, he calls the companies and finds out from them that the real average rent for all 40 apartments in his four blocks is \$90.00 per month. Given this information, set up as many problems as you can about Mark and George's survey.

APPENDIX F

CHILDHOOD ATTITUDE INVENTORY FOR PROBLEM SOLVING*

---

*Covington, M. V. Childhood Attitude Inventory For Problem Solving, Berkeley, California: University of California, 1967. Used by permission of the author.

Some engineers want to run a heavy television cable through a pipe. The pipe is 500 feet long and 6 inches in diamater. The pipe is about 10 feet below the ground, but it is open at both ends so the engineers can work on it. The pipe is not straight, but is made up of sections that twist and curve.

The engineers have already tried to push the television cable through from the ends of the pipe, but each time the cable twists and gets stuck after only a few feet.

The problem is to think of ways to run this particular television cable through this pipe, but without ripping up or digging down to the pipe.

DO NOT WRITE IN THIS BOOKLET.

DO NOT TURN THIS PAGE UNTIL YOU ARE TOLD TO DO SO.

PART I

Let's pretend that your class has been given this problem to solve. Here are some things college students might say about solving this problem, or problems like it. Circle, "Yes" if you agree with a statement, or "No" if you disagree. Remember, there are no right or wrong answers.

1. A problem like this one is probably too hard for anyone like myself to solve.
2. There is probably only one answer to a problem like this one.
3. If someone gets an idea that no one else has thought of, he should keep it to himself.
4. In a problem like this one, the best answer will be the one that most of the class decided is right.
5. If someone gets an idea that is different from everyone else's, the idea is probably not very good, otherwise others would have thought of it too.
6. Ideas that are wrong don't need to be suggested, because they only waste time.
7. Most students like social studies better than science.
8. When there is a hard problem to solve, it doesn't help very much to have someone in class who gets unusual ideas that no one else thinks of.
9. An idea for solving a problem that leads to a wrong answer still might be a good idea.
10. It would be best if everyone decided on one answer to this problem.
11. In most problems, poor ideas will not lead students to the right answer.
12. In solving problems, the most important thing to do is to try to figure out what is wrong with the ideas suggested rather than to think of new ideas.

13. The best answer should be the one that the teacher thinks is right.

14. New ideas should be tried out only after older ideas that have worked before fail to bring an answer.

15. In solving problems, one should consider the silly ideas as well as the ideas that have worked before.

16. Although several answers may be suggested, there is usually only one best answer.

17. Someone who keeps working on a problem that no one else in the class can solve is stubborn and selfish.

18. Problems are not fair if they can't be solved by using rules which everyone knows.

19. Anyone who asks questions after the teacher has finished giving directions is probably too lazy to solve the problem for himself.

20. Anyone who suggests a lot of ideas usually keeps others from giving their ideas.

21.. Problems are not fair if they make you keep looking for new ideas in order to solve them.

22. If someone is not very good at thinking and solving problems by the time he is my age, then it is too late.

23. The best workers will get one good idea and stick with it rather than think of many ideas, which might confuse them.

24. It is best to make sure than an idea is a good one before it is suggested to the class.

25. Some people are naturally born to be better thinkers than others, and there is nothing that can be done.

26. There is probably only one way that is best for solving a problem like this one.

27. If no one is able to solve a problem like this one after a few minutes, then the problem is too hard for the class to solve.

28. In a problem like this one, the best answer should be the one that most members of the class think is right.

29. On problems like these, it is best not to know too much; otherwise people will become confused.

30. Young people can learn to read and to arithmetic, but they cannot learn to think better or get better ideas.

PART II

Now we would like to know how you, yourself, might feel when working on the problem of getting the television cable through the pipe. Remember, these questions are about how you think and feel, so there are no right or wrong answers.

1. Do you feel that other members of your class know more about what to do in working on a problem like this one than you do?
2. Do you feel that the best thing about school is lunch hour?
3. Do you feel that you would know how to get started on a problem like this one?
4. Do you feel that you would be unable to solve problems like this one, even if you worked hard on them?
5. Would you like to work on a problem like this one?
6. If you worked on this problem, do you think that you would get any good ideas?
7. Do you think that many times your suggestions and ideas are not taken seriously by your friends?
8. Do you feel that you shouldn't ask too many questions about problems in class?
9. Do you think that ideas given by other students for solving this problem would be better than your ideas?
10. Do you enjoy drawing maps and pictures more than anything else in school?
11. Do you think that other students know more about problems of this kind than you do?
12. Do you feel that your ideas might be laughed at?
13. If you already had one good idea, would you rather stick with it than look for more ideas?

14. Do you feel that you would not have as good a chance of solving this kind of problem as other students would?

15. Would you like to work on a problem like this one, even though you might not be able to solve it?

16. Although others might not laugh out loud at your ideas, do you still feel that they would not like them?

17. Do you think that your ideas for solving this problem would be about as good as the ideas given by the other students?

18. If most of the others had solved this problem, but you had not, would you want to give up?

19. Do you feel that you are one of those people who is just not very good at thinking and solving problems?

20. Do you feel that others in your class would understand the problem better than you?

21. If you got an idea that no one else thought of, would you keep it to yourself, even if you were told to share ideas?

22. Would you want to give up after some of the other students got ideas and you didn't?

APPENDIX G

INTERCORRELATIONS OF PROBLEM SOLVING TESTS, IQ, ARITHMETIC PRE-TEST, AND CREATIVITY TOTAL SCORE

Correlation Matrix Based on Total Sample (N=370)[1]

	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
1. IQ	49	43	23	40	41	60	48	06	37	31	37	23	41	32	14	44	1.0
2. Arith. Pre.	24	38	09	48	48	85	54	15	41	28	42	37	47	48	26	1.00	
3. GPS-1	14	18	02	24	10	18	25	03	10	07	17	13	16	27	1.00		
4. GPS-2	33	26	-03	40	32	47	36	05	32	14	25	30	56	1.00			
5. GPS-3	35	29	04	39	35	48	40	-02	27	22	28	33	1.00				
6. GPS-4	13	21	03	27	33	39	25	10	25	24	22	1.00					
7. GPS-5	15	24	03	30	34	45	36	18	33	19	1.00						
8. GPS-6	23	26	14	23	26	33	30	09	21	1.00							
9. GPS-7	09	34	17	35	39	49	44	26	1.00								
10. GPS-8	-07	05	04	04	19	09	18	1.00									
11. APS-Puzzle	25	42	16	45	38	ϕ	1.00										
12. APS-Text	28	41	09	51	50	1.00											
13. Make Up Prob.	22	28	10	30	1.00												
14. Attitude I	25	85	25	1.00													
15. Attitude II	22	72	1.00														
16. Attitude Tot.	28	1.00															
17. Post. Tot. Creat.	1.00																

---

[1] N=180 for APS-Puzzle, N=179 for APS-Text.; All decimal points ommitted; all R's rounded to 2 decimal places.

APPENDIX H

SUMMARY OF CATEGORIES ASSIGNED TO PUPILS' PROGRAM RESPONSES

Summary of Category Assignments of Pupils' Program Responses

Categories:
1. Memory-Organization
2. Production
3. Reorganization
4. Judgment-Evaluation
5. Attitude

Lesson-Page	Category Assignment
2 - 9	2
12	2
17A	2
17B	2
24	2
29	4
31	3
33	3
35	4
36	4
3 - 7	5
13	1
17	2
19	4
25	4
29	1
31	2

Lesson-Page	Category	Lesson-Page	Category
4 - 7	2	7 - 7	1
13	5	8	2
15	3	9	2
17	2	10	2
23	2	11	2
25	3	12	3
29	3	13	2
31	2	15A	4
33	4	15B	4
35	4	16	5
5 - 10	4	17	5
15	2	18A-E (Each)	5
17	5	8 - 9	2
21	1	11	2
25A	2	13	2
25B	2	17	2
29	2	23	4
31	1	25	4
35	4	29	2
37	3	31	2
39	4	33	1
6 - 5	5	35	4
7	1	9 - 11	4
19	4	13A	2
27	1	13B	2
31	1	15	2

Lesson-Page	Category
9 - 17	2
20	1
21	2
23	4
25A	4
25B	3
27	1
31	3
33	2
37	2
39A	1
39B	1
41	2
43	1
10 - 19	2
21	2
23	3
25A	4
25B	3
27	3
31	1
33	4
35	2
37	2
41	3
43	4
11 - 11	2
15	3
17	2
19	4
23	3
25	2
29	4
31	3
33	4
35	2
37	3
39	4
43	4
12 - 7	1
9	5
11	1
13	1
15	1
17	5
37	4
39	3
43	4

Lesson-Page	Category
13 - 12	1
13	1
14	2
15	2
16	2
17	3
18	2
19A	2
19B	2
19C	2
20	4
21	2
23A	4
23B	4
24	5
25	5
26-A-E (Each)	5
14 - 11	2
15	3
17	2
21	3
23	2
25	2
27	3
33	4
14 - 37A	1
37B	3
39	4
15 - 9	2
13A	4
13B	4
13C	2
13D	2
15	2
23	2
25	2
27	4
31	4
35A	4
35B	2
39	4
16 - 3	2
5	1
13	4
14	4
19	4
20	1
21	4
23	4
27	4

Total By Categories

Category	Total Items
1	24
2	62
3	23
4	46
5	20
	175

APPENDIX I

TEACHER RATING PROCEDURE AND CREATIVITY DEFINITION

Teacher Rating of Pupil Creativity

I. INTRODUCTION

As a part of your work with the Elementary School Creativity Project, we would like to ask you to rate your pupils on creativity.

In order to do this as effectively as possible, we ask you to read the following pages carefully, and to follow the instructions that are given.

These procedures should not require more than a few minutes of your time each day. Making the final ranking on the cards (as described in the directions) should not require more than about half an hour.

Your cooperation in this matter will enable us to evaluate the project, and our tests, with considerably greater precision than would be possible otherwise.

II. DEFINITION OF CREATIVE BEHAVIOR

Part A presents several short descriptions of how creative students usually act, which should help you in your ranking of students. Part B presents a general description of the creative student's composition work. You will note that the kinds of behavior described can generally be observed best in the open-ended discussion type of class.

A. Creative students typically:

like to do their own planning, make their own decisions, and need the least training and experience in self-guidance.

do not like to work with others, and prefer their own judgments of their work to the judgments of others. They therefore seldom ask other students (or their teachers) for their opinions in this respect.

take a hopeful outlook when presented with complex, difficult tasks.

have the most ideas when a chance to express individual opinion is presented. These ideas frequently invoke the ridicule of the class.

are much more likely to stand their ground in the face of criticism.

are the most resourceful when unusual circumstances arise.

can tolerate uncertainty and ambiguity better than others.

are not necessarily the "smartest" or "best" students.

B. In their compositions, creative students typically:

show an imaginative use of many different words.

are more flexible, e.g., in a narrative, they use more situations, characters, and settings. Rather than taking one clearly defined train of thought and pursuing it to its logical conclusion, creative students tend to switch the main focus quickly and easily, and often go off on tangents.

tend to elaborate on the topic assigned, taking much broader connotations of it to begin with, and then proceed to embellish even that.

are more original. (This is the most important characteristic. The others need not be evidenced, but this one must be.) This student's ideas are simply different from the average student's response. Perhaps you might react to the creative student's work in this way: "I know what most of the kids will do with the topic, but I never know what to expect of this one!"

III. MAKING THE RANKING

We will provide you, at or before the final seminar, with a pack of cards on which your pupils' names are written.

Until then, familiarize yourself with the creativity definition;

Observe your pupils at least once each day, giving attention to the behavior characteristics in the definition.

When you receive your cards, we would like you to rank your children, using the following procedure:

1. Begin by dividing the cards into two groups on your desk. In the group at your left, place the cards of pupils who are like the definition of creativity. At your right, place the cards of those who are unlike the definition. (See figure 1.)

2. Then divide each group again. Now, the group at the extreme left represents those most like the definition. Moving from left to right, the other three groups represent: moderately like the definition; moderately unlike the definition; and, most unlike the definition. (See Figure 2.)

3. Then spread out the cards in each group vertically, so that the cards of the pupils most like the definition of any in that group are at the top of the pile. (See Figure 3.)

4. Assign a number to each student's card (writing the appropriate number lightly on the front of each card, after the pupil's name). The cards should be numbered consecutively, beginning at the top of the column at the extreme left, and working down from the top of each subsequent column. (See Figure 4.)

5. The result should be a rank ordering of your rating of the comparison of your pupils to the definition of creativity. The last number should be equal to the total number of pupils in your class. A ranking of "1" indicates that the pupil is rated as being very much like the definition of creativity; a rank of "20" or "21" (as examples) will identify a pupil rather unlike the definition.

6. Return the completed deck of cards to us in the enclosed envelope.

FIGURE ONE

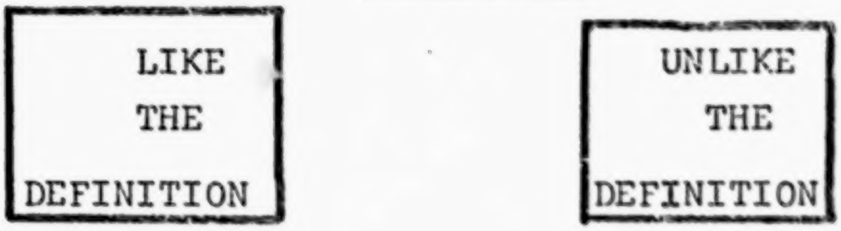


FIGURE TWO

MOST LIKE	MODERATELY LIKE	MODERATELY UNLIKE	MOST UNLIKE

FIGURE THREE

MOST
↑
LIKE THE DEFINITION
WITHIN EACH GROUP
↓
LEAST

FIGURE FOUR

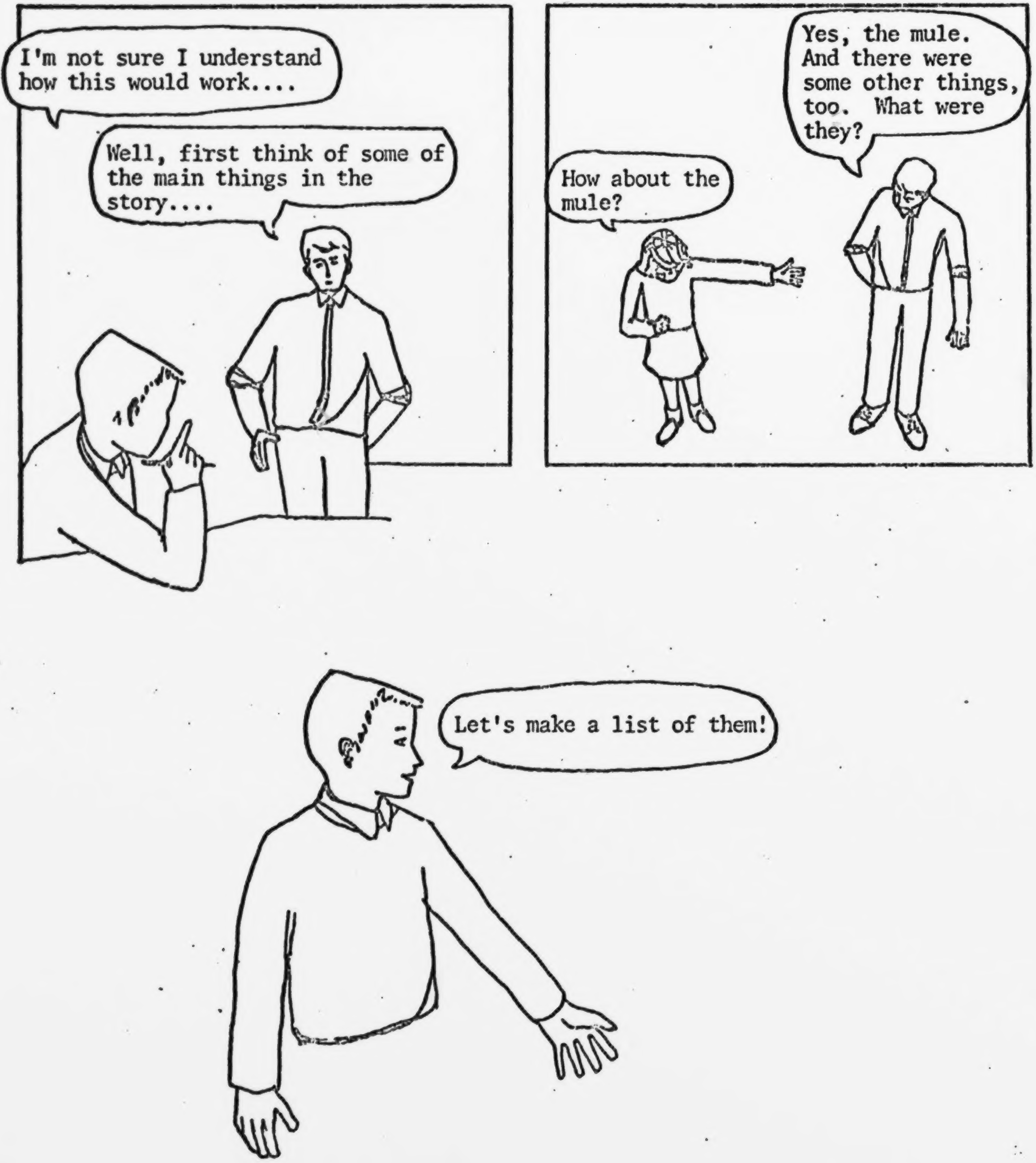
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
13.
14.
.
.
.

APPENDIX J

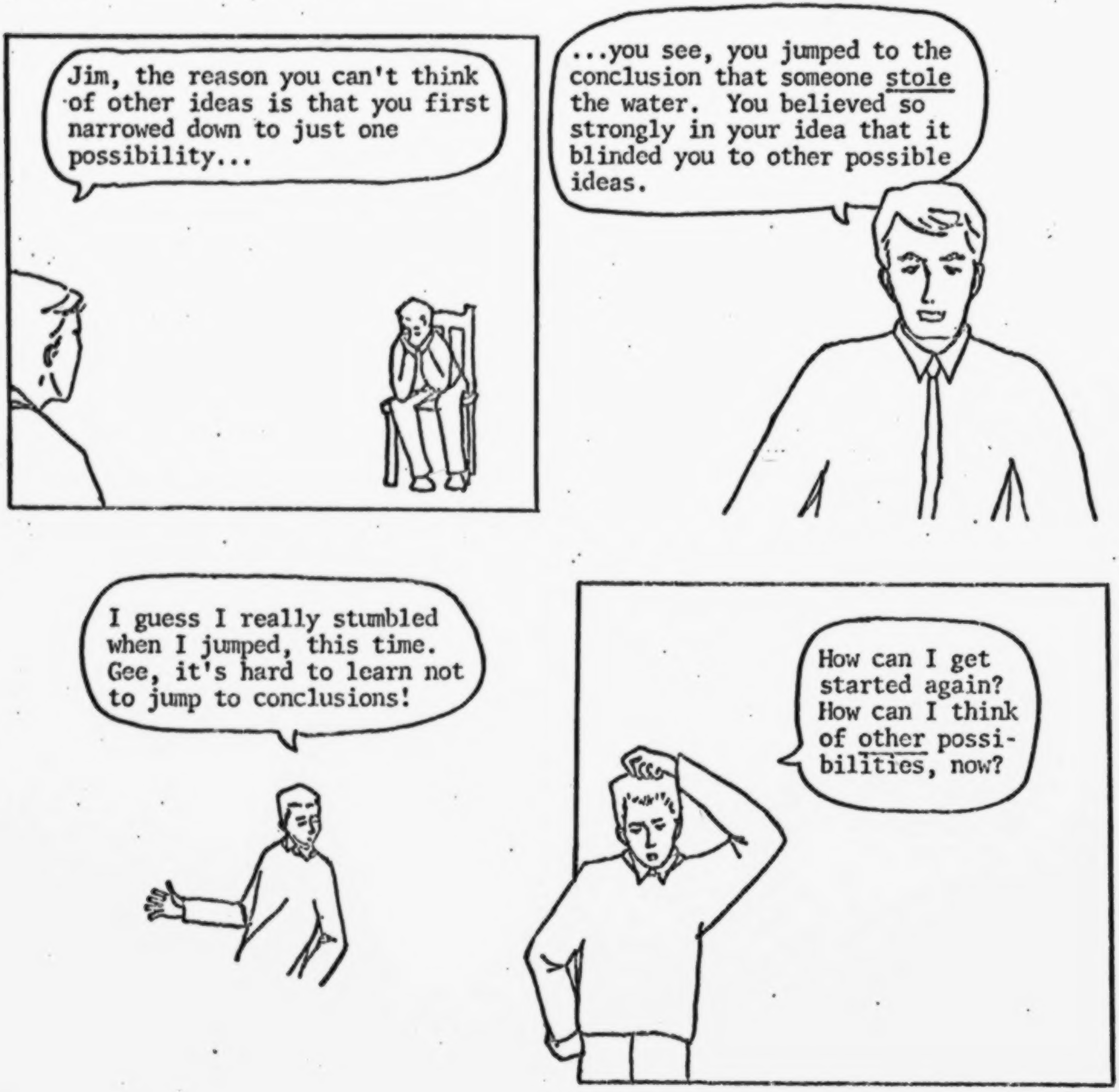
SAMPLE PAGES FROM THE PRODUCTIVE THINKING PROGRAM[1]

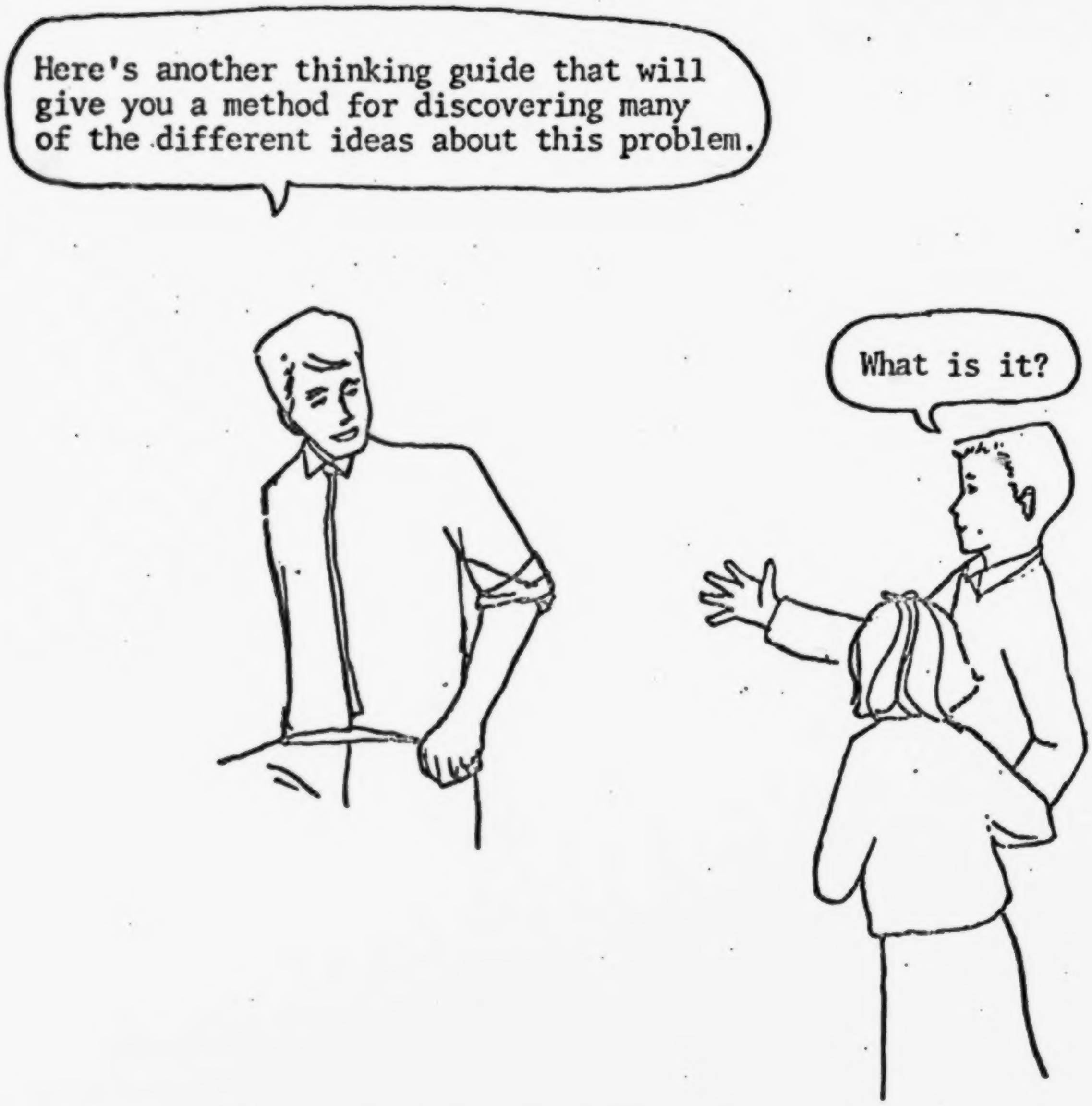
---

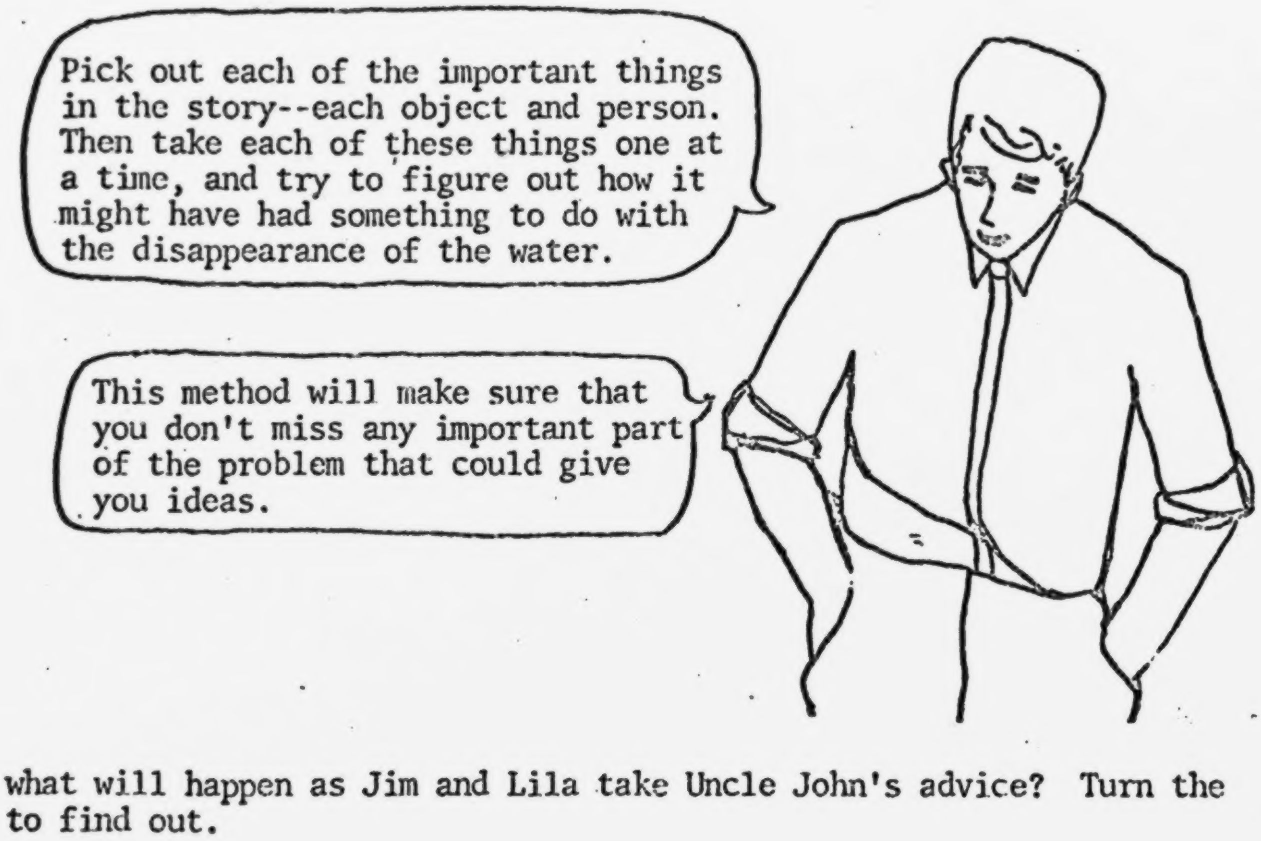
[1] Copyright 1966 by M. V. Covington, R. S. Crutchfield, and L. Davies. Booklet 5, pages 18-21; reproduced by permission of the authors.


I'm not sure I understand how this would work....
Well, first think of some of the main things in the story....
How about the mule?
Yes, the mule. And there were some other things, too. What were they?
Let's make a list of them!

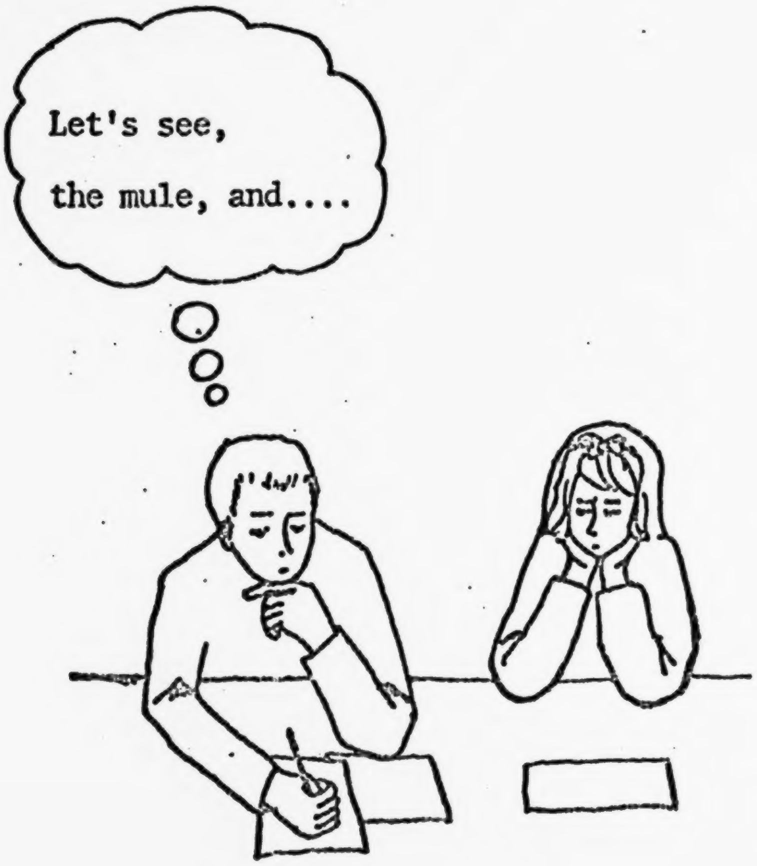
Uncle John notices Jim's silence:


Jim, the reason you can't think of other ideas is that you first narrowed down to just one possibility...
...you see, you jumped to the conclusion that someone stole the water. You believed so strongly in your idea that it blinded you to other possible ideas.
I guess I really stumbled when I jumped, this time. Gee, it's hard to learn not to jump to conclusions!
How can I get started again? How can I think of other possibilities, now?


Here's another thinking guide that will give you a method for discovering many of the different ideas about this problem.
What is it?


Pick out each of the important things in the story--each object and person. Then take each of these things one at a time, and try to figure out how it might have had something to do with the disappearance of the water.
This method will make sure that you don't miss any important part of the problem that could give you ideas.

Now, what will happen as Jim and Lila take Uncle John's advice? Turn the page to find out.


Let's see,
the mule, and....

You try making a list, too. Go back to pages 8 and 9 and read the story again. Then pick out each of the main things in the story and write it down:

____

____

____

____

____

____

____